2018

Multiple Functions Of The Striated Rootlet Proteins Of The Paramecium Basal Body

Md Ashikun Nabi
University of Vermont

MULTIPLE FUNCTIONS OF THE STRIATED ROOTLET PROTEINS OF THE *PARAMECIUM* BASAL BODY

A Dissertation Presented

by

Md Ashikun Nabi

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
Specializing in Biology

October, 2018

Defense Date: August 9, 2018
Dissertation Examination Committee:

Judith L. Van Houten Ph.D., Advisor
Alan K. Howe Ph.D., Chairperson
Bryan A. Ballif, Ph.D.
Alicia M. Ebert, Ph.D.
Cynthia J. Forehand Ph.D., Dean of the Graduate College

ABSTRACT

Paramecium ciliary basal bodies align in straight rows from posterior to anterior. Each basal body is connected to three rootlets ((Post Ciliary Rootlet (PCR), Transverse Rootlet (TR) and Striated Rootlet (SR)). The SR, the longest, projects from the basal body toward the anterior past several more anterior basal bodies. The depletion of Meckelin (MKS3) misaligns SRs, disorganizes basal body rows and makes the SRs appear ragged and serpentine. In this study we clarify the composition of the *Paramecium* ciliary basal body's SR and demonstrate that the SR plays a critical role in creating the orderly array of basal bodies in rows that run from pole to pole of the cell, likely through the interactions with centrins and other cytoskeletal elements underlying the cell surface. Here in this study we first report the reciprocal relationship between the SR and centrin related infraciliary lattice (ICL) protein that can dictate the cell surface morphology.

The SR of *Chlamydomonas* is the best studied. Using the single SR *Chlamydomonas* gene SF-assemblin to search in Paramecium DB, we found thirty *Paramecium* genes in thirteen Paralog Groups. Proteins from 13 paralog groups were confirmed to be in the SR structure using immunofluorescence. LC-MS/MS analyses of density fractions from SRs isolation show all thirty SR members are within the same density fraction. We further categorized all 30 *SR* genes in five Structural Groups based on their ability to form coiled coil domain and evaluate the function of all five Structural Group using RNA interference (RNAi). Silencing the transcripts of the any of the Structural Group showed misaligned basal body rows and the disordered organization of the SRs with abnormal appearance of SRs all over the cell surface. Silencing of Paralog Group showed normal phenotype except for the two Paralog Group (Paralog Group 1 or Paralog Group 7) which themselves constitute Structural Group individually. Isolated SRs from the control or Paralog Group depleted cells show a characteristic striation pattern that includes characteristic major and minor striations. Isolated SRs from any of the Structural Group depleted cells demonstrate abnormal shapes and striation periodicity. There is a correlation between the SR Structural Group RNAi surface misalignment phenotype and the isolated SR Structural Group RNAi phenotype for shape and periodicity of the SR. Strikingly our study of SR clearly demonstrates the role of SRs in shaping the other cytoskeleton structures of the cell cortex e.g., ICL, epiplasm territory and cortical unit territory.

In another follow up study of MKS3 (Picariello et al., 2014), we depleted the transcripts of *MKS5* gene in *Paramecium tetraurelia*. Depletion of MKS5 transcripts in *Paramecium* causes cilia loss all over the cell surface. Unlike MKS3 depletion, MKS5 depletion does not affect the straight basal body rows and the ordered organization of SRs. Moreover, data presented in this study clearly demonstrates depletion of MKS5 transcripts somehow affect the localization of another transition zone protein, B9D2.

It appears when lacking any of the SR Structural Group, the rest fail to interact properly with each other to maintain the SRs structure and directionality toward the anterior. As a result, abnormal SRs appear to lose the interaction with other cytoskeleton structures such as ICL network complex, which eventually results in misaligned basal body rows and altered swimming behavior. From the data presented in this study it is reasonable to postulate ICL1e subfamily and SRs are in a reciprocal relationship to maintain the straight basal body rows and the highly ordered organization of the SRs all over the cell surface.

ACKNOWLEDGEMENTS

I would first like to thank my advisor Dr. Judith Van Houten for providing me with the opportunity to work in her lab. Her constant encouragement and dedication helped make this possible. I would also like to thank Dr. Bryan Ballif, Dr. Alicia Ebert and Dr. Alan Howe for agreeing to be part of my dissertation committee. Your insight and critiques about this work were invaluable. I would like to thank Dr. Junji Yano and Megan Valentine for their amazing technical skills, suggestions and in depth discussion about this work. Thanks to all past members of the Van Houten lab, it was a pleasure to work with you.

Thanks to Dr. Ying Wai Lam, Dr. Bin Deng and Julia Fields for their expert technical assistance with the mass spectrometry analysis. Thanks to Todd Clason for his patience and his technical assistance with my immunofluorescence imaging. Thanks to Dr. Janine Beisson for the gift of the anti-SR antibody and the opportunity to use it in this project. Continued thanks to Anne Aubusson-Fleury for the anti-CTS32 antibody for epiplasm staining and Dr. Mark Winey for his gift of the anti-centrin antibody. Without either of these the project would not have been possible.

Finally, a special thank you to my parents who have supported me through my entire life. I want to extend my thanks to my wife, Nahin. Without all of you none of this would be possible.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 2

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 1: Comprehensive Literature Review

I. General overview:

Paramecium tetraurelia is a free living, unicellular eukaryotic organism. It is widely distributed in the fresh water habitat where it feeds on bacteria, yeast and algae. The size of the organism ranges from ~100 - 250 µm. It is often studied as a representative member of the ciliate group and has been widely used in class rooms and laboratories to study various biological processes. *Paramecium* is an excellent model for studying the cilia especially the biology of the basal body, centrioles and ciliopathy genes (Fig.1.1). The characteristics that make *Paramecium* an attractive model include high number of basal bodies (~4000 basal bodies per cell) and cilia, the sequenced and annotated genome, the presence of mammalian orthologs of ciliopathy genes and finally, the ease of applying various genetic and biochemical approaches (Vincensini et al., 2011). Like all other ciliates, the *Paramecium* cell has two distinct types of nuclei, a macronucleus and one or more micronuclei. The macronucleus is polyploid and contains 39,642 protein coding genes. The macronucleus is responsible for normal cellular function that controls non-reproductive cell functions, expressing the genes needed for daily functioning. The micronucleus is diploid, containing the genetic material that is passed along from one generation to the next. *Paramecium* reproduces asexually by binary fission. During asexual reproduction, the macronucleus divides by a process generally known as amitosis and the micronuclei divide by mitosis. Eventually the cell divides transversally and each new cell acquires a copy of the macronucleus and the micronucleus. Under certain conditions, self-fertilization (autogamy) may happen to *Paramecium.* The sexual reproduction method in *Paramecium* is known as conjugation in

which compatible mating types fuse temporarily to exchange genetic material. During conjugation, meiosis occurs to the micronuclei of each conjugant which results in haploid gametes that pass from one cell to the other. Diploid micronuclei form by the fusion of gametes from each conjugate. New macronuclei in each conjugant develop from the newly formed diploid micronuclei and the old macronuclei are destroyed. Sexual reproduction is triggered by starvation and therefore can be induced or prevented by controlling the availability of food. The life cycle of *Paramecium* is described in Fig.1.2.

Motility and behavior of *Paramecium tetraurelia*:

Paramecium cells move (forward or reverse) by the coordinate movement of thousands of cilia that cove the cell surface. Each cilium goes through a beat cycle consisting of power (or effective) stroke and recovery (or return) stroke. In a power stroke, which allows the cell to move forward, the cilium makes obliquely backward movement to the longitudinal axis of the cell body with force against the surrounding medium. The power stroke is followed by a return stroke. In a return stroke, the cilium makes a counter-clockwise rotation parallel to the cell surface for returning to its regional curved position (Kung et al., 1975; Satir et al., 2014).

Changes in behavior of *Paramecium* cells in response to different environmental stimuli (such as different chemicals, heat, touch and light) have been described in many previous studies (Van Houten, 1998; Eckert et al. 1972). Although cell movement is controlled by the movement of the cilia, the ciliary activity is dictated exclusively by the membrane potential of the cilia. The cell activates cilia associated voltage sensitive Ca^{2+} current (I_{ca}) when it encounters depolarizing stimuli (chemicals) or collides with an

object. If the stimulus is sufficiently strong, it results in an intraciliary Ca^{2+} concentration high enough to cause the cilia to beat in reverse direction, which eventually results in backward swimming of the cell. The return of the cell to its resting membrane potential (-30 to -40 mV) causes the power stroke to return to normal. The cell accomplishes this through the action of two potassium currents, the voltage activated K_v(Satow and Kung, 1980) and the Ca^{2+} activated K_{Ca} (Brehm et al., 1978). The cell will swim forward when the ciliary Ca^{2+} is reduced again to nM levels (Eckert, 1972). The voltage gated Ca^{2+} channels (Ca_V) that are exclusively found in cilia play an important role in the reversal of ciliary beating (Lodh et al., 2016) .

Cilia and basal bodies:

The basal body is the microtubule-based symmetrical structure that anchors the cilium. The core structure of the basal body consists of nine sets of microtubule triplets that align in a cylindrical shape. The mature and functional architecture of the basal body shows several distinct regions. The proximal end of the basal body exhibits a characteristic cartwheel structure, the middle lumen region that lacks the cartwheel structure and the distal end is characterized by the transition fiber. In the proximal end, in the cartwheel structure**,** the A- and C-tubules of adjacent triplet microtubules are connected by an A-C linkage. A central hub with nine spokes connects to the A-tubule of each triplet microtubule blade. An electron dense collar structure surrounds the proximal end of the basal body and serves as the attachment site to the cortical structure (Frankel 2000). Many proteins interact in these functionally distinct regions and play vital roles in the maintenance of the organization of the mature basal body, assembly of new basal

bodies, basal body cortical attachment, ciliary assembly and circumferential polarity (Kilburn et al., 2007).

Cilia are hair-like projections extending from the basal body on the cell surface. The structure and the organization of the cilia have distinct features as compared to other organelles in the cell. Although the ciliary membrane is continuous with the plasma membrane of the cell, the constituent compositions of these two membranes differ remarkably from each other (Rohatgi and Snell, 2010). The nine-fold rotational microtubule triplets of the basal body structure organize the microtubule doublets of ciliary axoneme structure. Unlike the basal body, the ciliary axoneme contains 9 outer doublets of microtubules that are extensions of the A and B tubules of the basal body (Ringo, 1967). In addition to the nine outer doublet microtubules, cilia can be characterized by the presence or absence of a central pair of microtubules: 9+2 structure and the 9+0 structure. Cilia function as the cell's sensory antenna regardless of the presence or absence of the central pair of microtubules (Bloodgood, 2010; Christensen et al., 2007). Ciliary structures found in *Paramecium* are the classic example of the 9+2 structure of the cilia (Fig.1.3).

The transition zone (TZ) is localized just above the plane of the basal body in a region marked by characteristic Y linkers (Szymanska and Johnson, 2012). In the TZ region, proteins interact to form a molecular gate that controls the entry and exit of various proteins in the ciliary axoneme compartment. The TZ is also the region where basal body microtubule triplets convert to microtubule doublets structure to organize the core of ciliary axoneme (Czarnecki and Shah, 2012; Kobayashi and Dynlacht, 2011). A number of human ciliopathy proteins localize at the TZ of the ciliary compartment.

Multiprotein complexes that interact at the TZ include nephronophthisis (NPHP) 1-4-8, Meckel-Gruber syndrome (MKS), NPHP 5-6 and inversion compartment complex. Different members of MKS complex are MKS1, MKS2 (TMEM216), MKS3 (Meckelin/TMEM67), MKS5 (RPGRIP1L), MKS6 (CC2D2A), MKS8 (TCTN2), MKS9 (B9D1) and MKS10 (B9D2). MKS5 protein localizes independently and influences the organization of MKS-B9 multiprotein complex at the TZ by interacting with other members of the complex (Czarnecki and Shah, 2012; Kobayashi and Dynlacht, 2011). The NPHP module of the TZ is composed of NPHP/nephrocystin-1, NPHP-4 and MKS4 and believed to take part in the regulation of intraflagellar transport (IFT) and cargo entry into the cilium. The MKS module contains MKS1, MKS3, MKS6 and MKS8. This module functions in neural tube development and Hedgehog signaling (Sang et al., 2011). Further studies have elucidated other interacting proteins in the TZ including the Tectonic module consisting of TCTN1, TCTN2 and TCTN3 and several more MKS module components including B9D1, B9D2 and MKS3. The tectonic module is believed to function in regulating ciliogenesis and ciliary membrane composition. The MKS module is linked to the NPHP module through MKS5 protein. Both modules function in early ciliogenesis including docking and transition zone formation (Garcia-Gonzalo et al., 2011; Williams et al., 2011). The TZ proteins that are found in *Paramecium* by BLAST searching in Paramecium DB are described in Fig.1.4.

Basal body duplication in *Paramecium*:

As in other ciliates, the *Paramecium* basal body is associated with a set of three rootlets: the Post Ciliary Rootlet (PCR), Transverse Rootlet (TR) and Striated Rootlet (SR). Duplication of the daughter basal bodies is a complex and tightly controlled process

in *Paramecium.* The mother basal body is the structure that nucleates the assembly of the daughter basal body. The daughter basal body grows and matures in close proximity to the mother basal body. As the daughter basal body matures, it moves away from the mother basal body and becomes parallel to the mother basal body. Eventually, the daughter basal body develops all the rootlets and forms the mature basal body with the new cortical unit. Localization of the daughter basal body relative to the mother basal body is an intriguing process. The molecular mechanism of this process is not fully understood. Early ultra-structure studies suggest that during duplication, localization of the daughter basal body is governed by its surrounding environment with the parental SR playing a scaffolding role in guiding the movement of the daughter basal body towards the anterior pole of the cell (Iftode and Fleury-Aubusson, 2003). The molecular mechanism of how the SR guides the localization of daughter basal bodies is not fully understood.

Basal body duplication in *Paramecium* is a complex process controlled by spatial and temporal regulation over the cell surface. There are invariant regions (such as the extreme anterior and posterior of the cell) of the cell surface where duplication of basal body is not evident (Iftode et al., 1989). In contrast, the highest basal body duplication activity occurs around the midline of the cell, where there is elongation of the cell to prepare for the development of a new anterior and a new posterior daughter cell. In this region, basal body duplication takes place in two waves, the first of which has two rounds. The first round produces a new basal body in each cortical unit resulting in two or three basal bodies per cortical unit. The second round creates another basal body in each cortical unit along a narrow band around the midline of the cell on both the dorsal

and ventral surfaces (Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003). This second round results in cortical units with a third or fourth basal body. The second wave of duplication reconstitutes either single or paired basal body units in specific regions of the cell depending on their position on the cell surface (Iftode et al., 1989). The simplified diagram of the wave of the basal body duplication is described in Fig.1.5.

Cortex of the *Paramecium* and the associated structure:

The cortex of the *Paramecium* is a highly organized structure, which is composed of the plasma membrane and a number of cytoskeleton networks in close association (Fig.1.6). The outer surface is divided into a network of the hexagonal ridges where each is defined as a cortical unit. All the cortical units are organized in straight rows from the posterior pole towards the anterior pole of the cell. In interphase cells, each cortical unit can harbor single or pair of basal bodies (Allen 1971; Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003; Tassin et al., 2016). The plasma membrane that covers the outer surface of the cortical unit remains in close contact with a submembrane cytoskeleton structure known as epiplasm (Allen, 1971; Damaj et al., 2009; Aubusson-Fleury et al., 2013). In the single basal body cortical unit, the basal body has a set of three asymmetrical rootlets. One of the rootlets is SR that projects from the posterior pole of the cell towards the anterior pole of the cell and spans several cortical units. The other two rootlets are the microtubule ribbons called the PCR that projects towards the posterior pole of the cell and the TR that project laterally towards the adjacent rows of basal bodies. In the paired basal bodies unit, the posterior basal body has all the three asymmetrical rootlets and the anterior basal body only has the TR rootlet. The anterior basal body attaches to the SR rootlet of the posterior basal body. The basal bodies along

with their three asymmetrical rootlets align in a highly organized straight line from the posterior to the anterior pole of the cell both in dorsal and ventral surfaces (Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003). Rootlets attach to the basal bodies at precise locations relative to the microtubule triplets. The SR, PCR and TR rootlets attach to the microtubule triple 5, 9 and 4 of the basal body respectively (Fig.1.7) (Jerka-Dziadosz et al., 2013).

Several other structures and cytoskeletal elements are associated with each cortical unit. Below the membrane are the alveolar sacs that are membrane bound organelles responsible for calcium storage (Allen, 1971). In addition, each cortical unit contains a parasomal sac, which is believed to be a site of endo-, exo- and pinocytosis (Allen and Fok, 1980; Plattner and Kissmehl, 2003). Closely apposed to the alveolar sac membrane is a filamentous layer called the epiplasm (Fig.1.6). The epiplasm is divided into regular shaped territories where basal body units occupy the center of the territory. The epiplasm is continuous under the alveolar sacs and contacts the distal region of the basal body cylinder (Allen, 1971). In *Paramecium*, the epiplasm is composed of 51 proteins known as epiplasmins that are subdivided into 5 phylogenic groups (Aubusson-Fleury et al., 2013). The epiplasm is a permanent structure that may function to solidify the orientation and positioning of newly assembled elements during division (Aubusson-Fleury et al., 2013; Aubusson-Fleury et al., 2012; Collins et al., 1980).

The infraciliary lattice (ICL) is the inner most cytoskeleton structure of the highly organized cortex of the *Paramecium*. It forms contractile filamentous bundles of irregular shaped polygonal meshes that run around the proximal end of the basal bodies (Fig.1.6) (Allen, 1971; Garreau de Loubresse et al., 1988; Garreau de Loubresse et al., 1991;

Gogendeau et al., 2008; Aubusson-Fleury et al. 2017). ICL is made of centrins and centrin binding proteins. There are 35 centrin proteins and 3 long Sfi1P like centrin binding proteins, which together constitute the ICL network in *Paramecium* (Garreau de Loubresse et al., 1988; Gogendeau et al., 2008; Aubusson-Fleury et al. 2017). The centrin proteins of ICL network can be divided into 10 subfamilies e.g., ICL1a, ICL1e, ICL3a, ICL3b, ICL5, ICL7, ICL8, ICL9, ICL10 and ICL11 (Gogendeau et al., 2008; Aubusson-Fleury et al. 2017). All the subfamilies of centrin are found in the ICL network and have specific functions related to the ICL network except the subfamily ICL1e. The ICL1e subfamily consists of seven ICL-centrin proteins (PtCen_icl1e, PtCen_icl1g, PtCen8, PtCen10, PtCen12, PtCen15 and PtCen18) that appear to have multiple locations in the cell including the ICL network (Gogendeau et al., 2008; Aubusson-Fleury et al., 2017). RNAi reduction of transcripts for this family suggests ICL1e sub family demonstrates that ICL1e sub family may function as physical link to attach the ICL network to the other cytoskeleton structure of the cell cortex (Gogendeau et al., 2008; Aubusson-Fleury et al., 2017).

SRs are the prominent cytoskeletal elements that are attached to basal body structures in the cortical units. The cytoskeleton structures with striations similar to SR are present in different systems such as ciliates (Rubin and Cunningham, 1973), amoeboflagellate (Larson and Dingle, 1981) and different epithelial ciliated cells (Lemullois et al., 1991). The SR in *Chlamydomonas* is commonly known as the striated microtubule-associated fiber (System I fiber) and is the best studied SR structure. The SR in green algae comprises a single polypeptide known as SF-assemblin. It is a 34 kDa protein (284 amino acid residues), has a N-terminal short head domain (31 amino acid

residues) and C-terminal helical rod domain which has the ability to form a segmented coiled coil structure (Hyams and King, 1985; Lechtreck and Melkonian, 1991; Patel et al., 1992; Weber et al., 1993;). In *Chlamydomonas*, both the N-terminal head and C-terminal rod domains of SF-assemblin protein are required for assembly and formation of the *Chlamydomonas* SR. Various deletion and insertions in both the N-terminal head domain and the C-terminal rod coiled coil domain changes the solubility, striation pattern or overall appearance of the SR (Lechtreck, 1998; Lechtreck et al 2002). In *Paramecium*, the SR projects from the basal body toward the anterior of the cell past several more anterior basal bodies. The SRs are long, slightly bent fibers in which tapering starts at approximately 700 nm. The length and thickness of the rootlets are variable depending on their location on the cell surface. The structure comprises proteins ranging from 30-36 kDa molecular mass (Sperling, 1989; Sperling et al., 1991). They interact to form a very long striated structure that is dynamic (i.e. changes length during the cell cycle) (Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003).

Intraflagellar transport (IFT):

As mentioned previously, the structure and the organization of the cilia have distinct features as compared to other organelles. This compartment of the cell is surrounded by the lipid bi-layer membrane, known as the ciliary membrane. Despite this connection to the plasma membrane, the protein and lipid composition of the ciliary membrane is distinct from the protein composition of the cell membrane (Rosenbaum and Witman, 2002; Rohatgi and Snell, 2010). As the cilia lack protein synthesis machinery, an exclusive bidirectional transport process known as intraflagellar transport (IFT) is required for the construction, maintenance and the function of cilia (Fig.1.8). IFT was

first identified in the green algae *Chlamydomonas reinhardtii* (Kozminski et al., 1993). IFT initiates with the assembly of its particles and cargo around the base of the cilium. Anterograde movement of the IFT particles with cargo is facilitated by kinesin motors from the base to the tip of the cilia. Retrograde movement of the IFT particles and cargo from the tip to the base of the cilia is powered by dynein motors. IFT particles are required for the axoneme growth and maintenance. IFT particles deliver the required components for the axonemal growth. As described before, cilia lack protein synthesis machinery so the axoneme always requires a supply of proteins. If the IFT particles are defective, the ciliary axoneme structure shrinks slowly. In case of functional ciliary axoneme, IFT particles are believed to carry used or turnover proteins from the ciliary tip to the base (Pazour et al., 1999; Sloboda, 2005; Davis et al., 2006; Pedersen et al., 2008; Carvalho-Santos et al., 2011). The trafficking of cilia specific proteins may be mediated by the ciliary target sequence or ciliary target motif (Emmer et al., 2010).

Most of the IFT particles have been characterized studying the model organism *Chlamydomonas reinhardtii.* There are mainly two sub-complexes that comprise IFT system. There are six members in the IFT A sub-complex which includes IFTs 43, 122A, 122B, 139, 140 and 144 (Cole et al., 1998; Piperno and Mead, 1997). The members of IFTA sub-complex primarily mediate the retrograde movement of cargo loaded with used components (turn over proteins). It occurs in the direction from the tip of the cilium to the base of the cilium. There are thirteen members in the IFT B sub-complex that includes the IFTs 20, 22, 25, 27, 46, 52, 57, 72, 74, 80, 81, 88 and 172 (Hao and Scholey, 2009; Lucker et al., 2010). The members of the IFT B sub-complex primarily are responsible for the delivery of cargos containing fresh supply of proteins required for proper ciliary

function. In general, defects in the members of IFT A sub-complex lead to mild defects in ciliary function but defects in the members of IFT B sub-complex lead to more severe defects in the function of cilia (Blacque et al., 2006; Hao and Scholey, 2009; Tsao and Gorovsky, 2008; Follit et al., 2009; Pazour et al., 2002; Pedersen and Rosenbaum, 2008).

Among the different IFT B sub-complex components, IFT 88 has been studied extensively (Pazour et al., 2000; Robert et al., 2007; Lucker et al., 2010). IFT 88 plays an important in the process of ciliogenesis and may also play an important role as regulator of the cell cycle. It is thought to be interacting with IFT46 and IFT52 at the core of the IFT B sub-complex and thought to mediate all the necessary interactions required to carry different cargo components (Robert et al., 2007; Lucker et al., 2010).

Ciliopathies:

Ciliopathies is the term used to describe the group of diseases that arise from the defects in the structure of the cilia or the dysfunction of the cilia. It is an expanding group of autosomal or X-linked disorders. In general, ciliopathies can be characterized by a number of symptoms, including multi-cystic kidneys, central nervous system (CNS) malformations, polydactyly, respiratory illness, liver malformations, retinal degeneration, diabetes, obesity and skeletal dysplasias (Afzelius, 2004; Badano et al.,2005; Badano et al., 2006). The study of the ciliary proteome (Gherman et al., 2006) reported more than 1000 proteins comprising various ion channels, molecular motors, IFT particles, signaling molecules, receptors and many more proteins. Among these proteins, mutations of over 40 genes and their associated proteins have been reported to contribute to ciliopathy diseases. Many of the gene mutations involved in ciliopathies encode proteins that localize to and function in basal bodies and the TZ (Waters and Beales, 2011).

Ciliopathy diseases in human can be heterogeneous phenotypically. Mutations in one single locus can contribute to different ciliopathy features although mutations in different loci can contribute to the same feature (Waters and Beales, 2011). Cilia and ciliary proteins can perform diverse functions in a tissue- and cell-specific manner. The expression of IFT 20 with some other IFT components (IFT57 and IFT88) in hematopoietic cells that apparently lack primary cilia and their role in the immune synapse is one of the examples of the diverse functions of ciliary proteins (Finetti et al., 2011). Therefore, investigation of the diverse functions of the ciliopathy proteins in the tissue- or cell-specific manner is the key to understand the heterogeneous nature of the ciliopathy diseases. Polycystic kidney disease, Bardet-Beidl syndrome, Joubert syndrome and Meckel syndrome are the example of some common ciliopathy diseases (Huang et al., 2007; Yu et al., 2007).

Polycystic kidney disease (PKD) is one of the most devastating classes of ciliopathy diseases demonstrated by the formation of kidney cysts. There are two forms of PKD. One is autosomal recessive polycystic kidney disease (ARPKD) characterized by the dysfunction of *PKHD1* gene. Another form of the disease is autosomal dominant polycystic kidney disease (ADPKD) characterized by the dysfunction of the *PKD1* and *PKD2* genes (Huang et al., 2007; Yu et al., 2007; Katsanis et al., 2001; Sharma et al., 2008; Nachury et al., 2007;).

Bardet-Biedl syndrome (BBS) is autosomal recessive ciliopathy disease characterized by obesity, kidney dysfunction, polydactyly and learning disabilities. There are 14 human genes which have been reported to be associated with BBS. Most of the proteins that are associated with BBS, are involved in trafficking and sorting of the

proteins targeted for the cilia (Blacque and Leroux, 2006; Katsanis et al., 2001; Sharma et al., 2008; Nachury et al., 2007; Valentine et al., 2012). Joubert syndrome and related disorders (JSRD is an autosomal recessive disorder) are characterized by polydactyly, obesity, renal failure and cerebellar hypoplasia (Braddock et al., 2007).

Common genetic and clinical aspects have been reported among some closely related ciliopathy diseases such as JSRD, Meckel Gruber syndrome (MKS), and Nephronophthisis. Different ciliopathy genes related to MKS include MKS1, MKS3, CEP290, MKS5, MKS6 and MKS2. Among these genes, MKS3, CEP290, MKS5, MKS6 and MKS2 are all allelic at several loci for both MKS and JSRD (Smith et al., 2006; Kyttala et al., 2006; Tallila et al., 2008; Valente et al., 2006). The most common clinical features of MKS include renal dysplasia, occipital encephalocele and polydactyly (Salonen and Norio1984; Salonen and Paavola, 1998). Mutations in different regions of the MKS3 protein have been found in a number of studies. Mutations of different types affect different individuals in a number of ways. Mutational analysis of the patients from different studies revealed 33 distinct mutations in the 28 exons of the TMEM67 gene except exons 4 and 28. Different mutations include missense, truncating, splice-site and rare multiexon deletions. In general, eighteen missense, twelve truncating (including frame-shift and nonsense) and three 5′ or 3′ canonical splice sites mutations have been described. The missense mutations in the region of exon 8 to 15 are associated with lethal or more severe phenotypes as compared to mutations in other regions of the exons, suggesting that region of the Meckelin polypeptide may have some important function (Smith et al., 2006). Localization of the MKS3 protein along with its multiprotein complex (MKS-B9) in the TZ has an essential role in ciliogenesis and has been reported

by a number of researchers (Dawe et al., 2007; Williams et al., 2008; Bialas et al., 2009; Cui et al., 2011). In addition, the MKS3 protein is involved in the maintenance of the length of cilia in different animal and cell line models, regulation of centrosome (analogous to the basal body structure) number, migration of centrioles to the apical surface of the cell and interaction with actin binding protein nesprin 2 and filamin A. (Smith et al. 2006; Dawe et al., 2007; Dawe et al., 2009; Adams et al., 2012). Recently the work of Abdelhamed et al. (2015) reported the function of the MKS3 protein as a receptor and its role in positioning the basal body and epithelial branching morphogenesis via non-canonical Wnt pathway.

II. Experimental approaches:

RNA interference:

RNA interference (RNAi), also known as post-transcriptional gene silencing, is a biological tool in which gene expression or translation of the targeted protein is inhibited by the neutralization of the mRNA. It was first described in the worm *Cenorhabditis elegans* (Fire et al., 1998). Since its introduction it has widely been used to investigate the roles of proteins in the cellular pathways. RNAi uses cell′s intrinsic machinery to neutralize the target mRNA. It is a RNA-dependent mechanism of the cell controlled by the RNA-induced silencing complex (RISC). The entire process is initiated by the double-stranded RNA molecule (ds-RNA) inside the cell. The ds-RNA molecules activate an RNase-III like enzyme complex known as DICER. The DICER complex binds to ds-RNA and cleaves it into short double-stranded fragments of ~21-25 nucleotide pieces known as small interfering RNAs (siRNAs). Each siRNA fragment then unwinds to form the guide strand and the passenger strand. The passenger strand

eventually degrades and the guide strand incorporates into the RNA inducing silencing complex (RISC). In general, RISC also includes the protein Argonaute2. The guide RNA binds to the complementary mRNA then Argonaute2 (catalytic component of the RISC complex) carries out the cleavage of the mRNA. This whole process results in post-transcriptional gene silencing (Liu et al., 2004). RNAi is used to investigate the effect of protein depletion both in cell culture and living organisms. The specificity of RNAi depends on the complementary nature of siRNAs and the endogenous mRNA. Ideally only the targeted mRNA will be degraded, but cleavage of similar sequences in other mRNA transcripts can occur which is known as off-target effects of RNAi. Off-target analysis must be taken into account when designing RNAi constructs and interpreting data (Agrawal et al., 2003; Dykxhoorn and Lieberman, 2005).

There are several ways to perform RNAi in cell culture and model organisms, such as injection, transfection, soaking and feeding (Ketting and Plasterk, 2004). In *Paramecium*, feeding is the most common and preferred method of RNAi implementation because of the feeding behavior of *Paramecium* and the efficiency of the method (Galvani and Sperling, 2002). The feeding procedure involves the use of the RNase-III deficient cell (Ht115 *E. coli* DE3) line which allows the expression of large amounts of ds-RNA inside the bacteria. The target mRNA sequence is inserted in a T7 promoter containing L4440 plasmid vector. Once the Ht115 *E. coli* DE3 is transformed with vector plasmid, bacterial culture is induced with Isopropyl β-D thiogalactopyranoside (IPTG) to generate ds-RNA. Paramecia are allowed to feed on these ds-RNA containing bacteria. Endogenous RNAi machinery of the paramecia

utilizes ds-RNA to deplete the mRNA of the target gene. The efficiency of the RNAi feeding method is 70-80% depending on the target gene (Galvani and Sperling, 2002).

For all the RNAi experiments in the thesis I inserted the target sequence in the L4440 vector plasmid and transformed Ht115 *E. coli* DE3 bacteria with the plasmid of interest. I fed an induced bacterial culture (containing ds-RNA inside bacteria) to 51s *P. tetraurelia* cells to deplete the message for the target gene of interest. In all RNAi experiments, empty L4440 plasmid vector was fed to 51s *P. tetraurelia* cells serving as positive control. The phenotype of different RNAi cells or control cells was analyzed by immunofluorescence.

Expression of exogenous protein in *Paramecium:*

In the macronucleus of the *Paramecium* cells it is possible to introduce exogenous tagged versions of *Paramecium* protein by injection. To do that the target gene sequence needs to be inserted in a plasmid vector. The most common vector used for the expression of exogenous tagged version of target protein is pPXV. The pPXV vector has calmodulin promoters, *Tetrahymena* telomeres, multiple restriction endonuclease sites and a 3x-FLAG tag (5′-DYKDDDDK-3′) sequence. The calmodulin promoter and the *Tetrahymena* telomers of the pPXV vector offer high level expression of target exogenous gene and vector stability inside host cell respectively. The presence of a number of restriction endonuclease sites in the plasmid vector allows for the insertion of the target sequence in such a way so that the 3x-FLAG tag can be added either at the N- or C-terminus of the target sequence. In general, after injection of the linearized form of the plasmid vector containing the target gene sequence paramecia cells express

exogenous protein for 4 weeks (Valentine et al., 2012; Picariello et al., 2014). The cells then can be used for localization studies or biochemical studies.

For thesis, I used commercial anti-FLAG antibody to detect the localization of the FLAG-tagged exogenous version of the proteins of interest. The same antibody was used for various immunoblot experiments.

One pitfall of the exogenous expression of the target gene of interest is the loss of expression due to autogamy in *Paramecium*. Autogamy is a meiotic pathway in *Paramecium* that can be triggered by the starvation of the cells. It results in the break down and subsequent re-formation of the macronucleus (Berger, 1986; Godiska et al., 1987). During the process of autogamy loss of exogenous vector plasmid happens which eventually results in the loss of the epitope-tagged protein expression.

Isolation of the SRs:

As mentioned before SR is the largest rootlet of the *Paramecium* basal body. Isolation of the SRs from the cell cortex cytoskeleton was described by Sperling et al., 1991. For this thesis, I used the SR isolation protocol with modification to isolate SRs from different *SR* gene-depleted (message depletion was done by RNAi) cells or control cells. Different concentrations of (35%, 30%, 25%, 20%, 15% and 0% Optiprep in PHEM buffer) Optiprep were used to isolate SRs from the mixture of crude rootlets.

I analyzed the isolated SRs from different *Paramecium* cells by immunofluorescence, electron microscopy (Transmission Electron Micrograph of the SRs after negative staining) and LC/MS-MS analysis.

One pitfall of the SRs isolation procedure is the contamination of the preparation with the associated basal body units.

Experimental Approach for the SR project:

The unicellular ciliate organism *Paramecium* is an excellent model for studying cilia, especially the biology of the basal body, centrioles and ciliopathy genes. Work that had previously been done in our lab showed that loss of MKS3 in *Paramecium* led to global ciliary loss and basal body row disorganization on the dorsal surface of the cell (Picariello et al., 2014). Picariello et al., 2014 also showed the depletion of MKS3 causes a disordered organization of the SRs. In this study, we used bioinformatic analysis to identify 30 SR proteins belonging to 13 paralog groups based on the availability of the genome sequence in the ParameciumDB and with the help of proteomic analysis of the isolated SRs. Further analysis of the SR proteins for their ability to form coiled coil domain and functional studies of the SR proteins by using RNAi classified these 30 proteins into five Structural Groups. The phenotype (basal body row alignment and the SRs organization from the posterior pole towards the anterior pole of the cell) of the different *SR* gene depleted cells was analyzed by immunofluorescence. The importance of all five Structural Groups to form a normal SR of appropriate length and striation was done using TEM of the isolated SRs. The N-FLAG epitope tagging of different *SR* genes followed by immunofluorescence allowed us to determine the location of the SR protein. The data presented in this study suggest all five Structural Groups of SR proteins are required to form a SR with the appropriate length and striation. The data further suggest the SR of the basal body unit play an important role to maintain the straight basal body rows and also to maintain the current shape of other cytoskeletal structure of the cell

cortex. Together, all the data presented in this study help to better understand an indispensible interaction between the SR and the innermost cytoskeleton structure (ICL) of the cell cortex. We propose that centrin ICL1e subfamily of the infraciliary lattice interacts with the SR of the basal body unit and both these cytoskeleton structures work together to maintain the basal body units in straight cortical rows.

Fig. 1.1: Scanning Electron Micrograph (SEM) Of *Paramecium tetraurelia. Courtesy of J. Van Houten*

Panel A is a SEM of a ciliated cell where the surface of the cell is covered by thousands of cilia. Panel B is a SEM image of a deciliated cell where the highly organized cortical units are visible.

The top is from the 1985 Grass Calendar. Courtesy of J. Van Houten

A

B

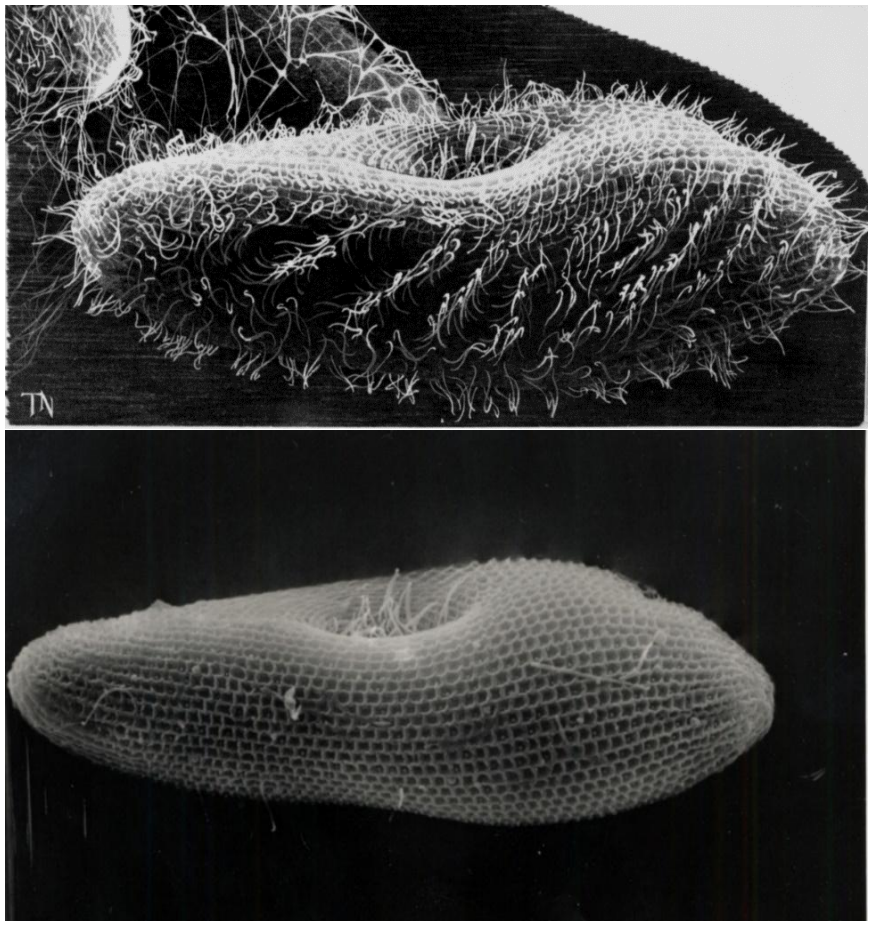


Fig. 1.2: Life cycle of *Paramecium tetraurelia.*

The sexual (conjugation) and asexual reproduction (self-fertilization and autogamy) life cycle in *Paramecium.*

Image Source: http://www.metamicrobe.com/paramecium/

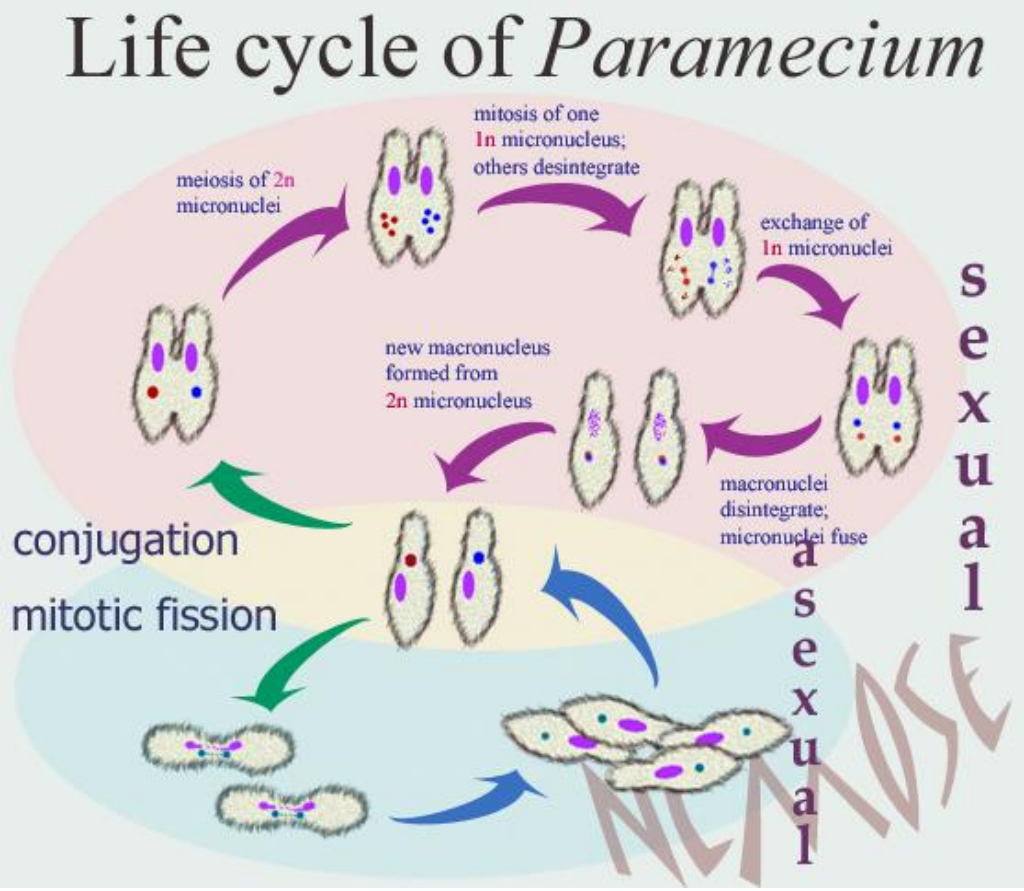


Fig. 1.3: Transverse section of the basal body structure.

The top panel shows the 9+2 and 9+0 axoneme arrangements. The middle panel shows the characteristic cartwheel structure at proximal and distal end of a basal body. A, B and C shows the microtubule triplets. The distal end of the basal body shows an electron dense material. Image modified from Dawe et al., 2006 with permission.

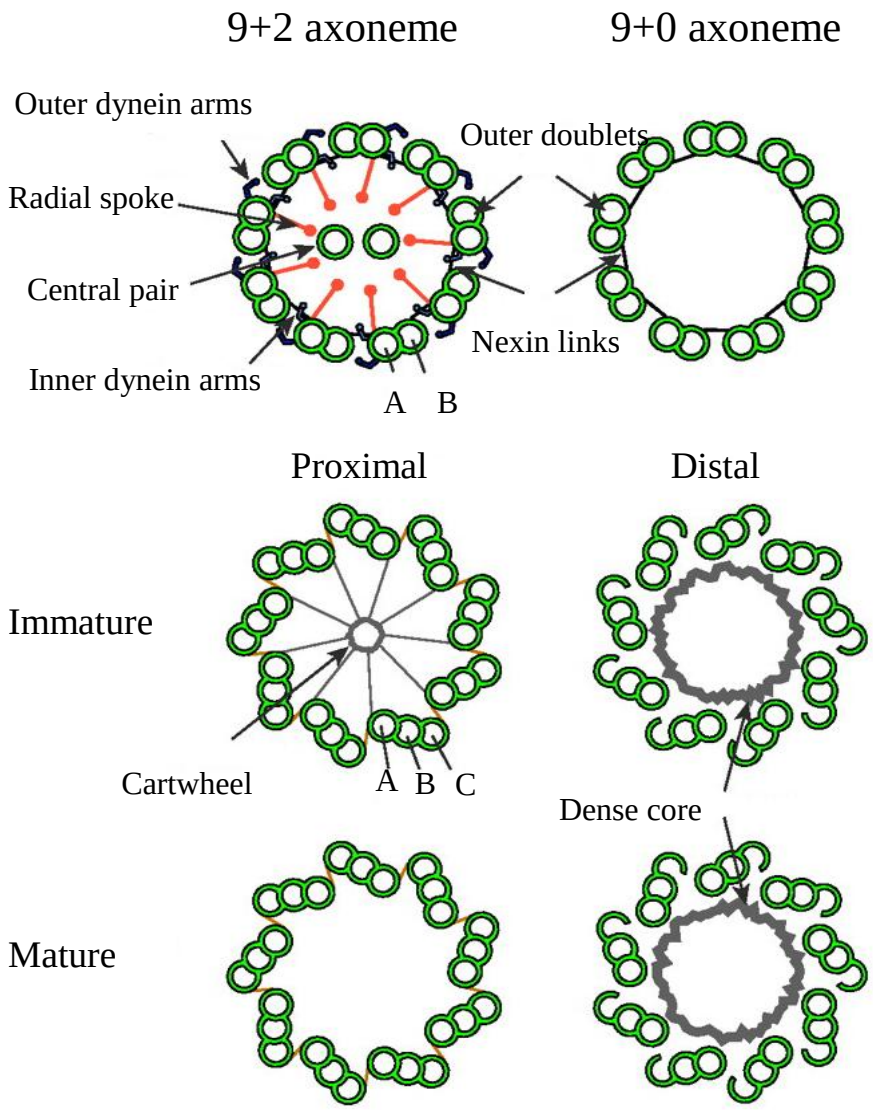


Fig. 1.4: Transition Zone (TZ) protein complexes in *Paramecium*.

Image is showing NPHP 1-4-8, MKS/B9, NPHP 5-6 and inversion compartment protein complexes in Paramecium identified by BLAST search in ParameciumDB. These protein complexes interact to each other to form a molecular gate that controls the entry and exit of proteins in ciliary structure.

Image is modified from Czarnecki and Shah, 2012 with permission.

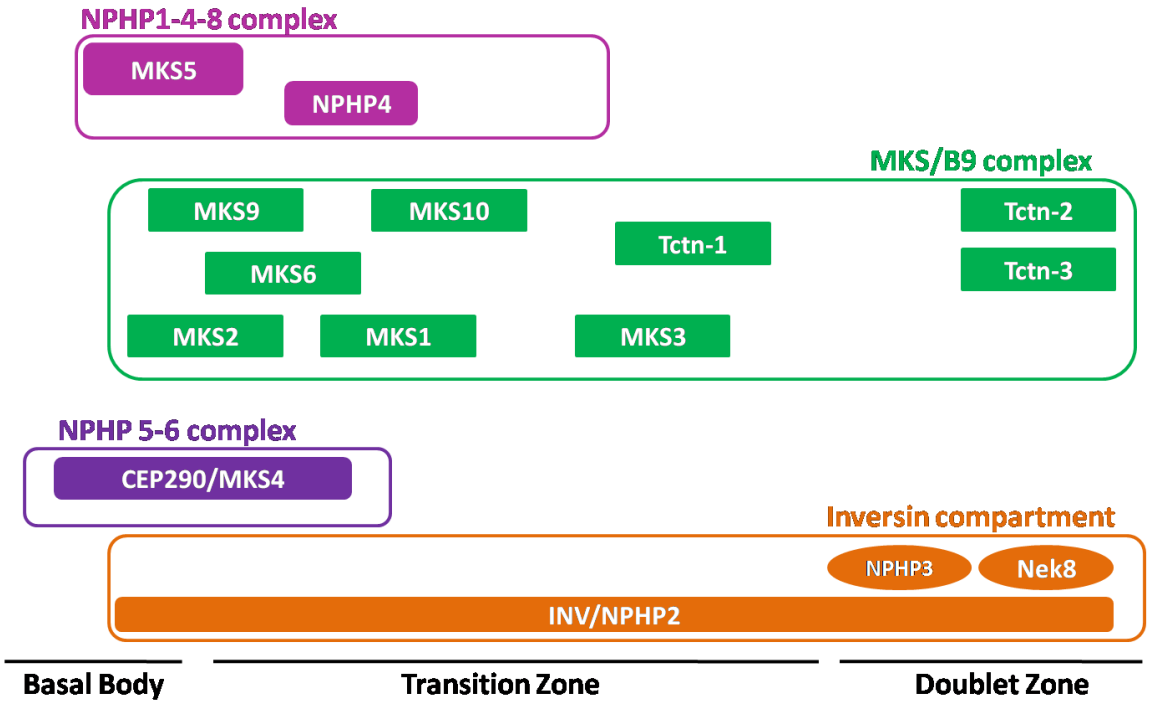


Fig. 1.5: Progression of the basal body duplication waves across the dorsal surface of *Paramecium.*

The first wave of basal body duplication occurs via two rounds of basal body duplication, highlighted in grey color in the image. It progresses outwards from the area of the fission furrow towards the anterior and posterior pole of the cell. Each round adds one basal body to duplicating cortical units. A second wave of basal body duplication starts near the fission furrow (demonstrated in black) as the first round of basal body duplication nears completion. The second wave of basal body duplication remains localized to distinct regions in the anterior and posterior daughter cells. The second wave reconstitutes two basal body units in these specific areas as well as in certain units across the dorsal surface.

Image is used from Iftode et al., 1989 with permission.

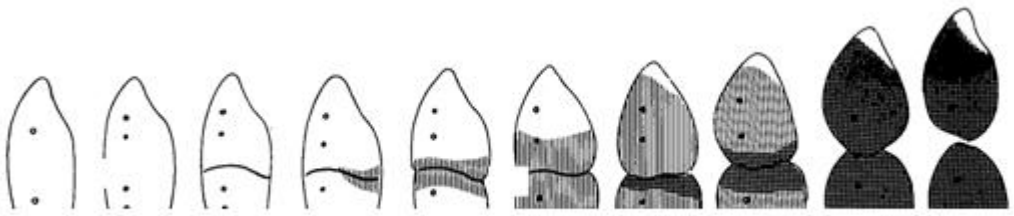


Fig. 1.6: The highly organized cortex structure of the *Paramecium*.

In Panel the cell cortex of *P. tetraurelia* shows several layers of organization of cytoskeleton structures. The outer surface of the cell is divided into organized cortical units. Each cortical unit harbors one or two basal body units from which cilia arise. Just below the cell membrane, a submembranous cytoskeleton layer defined as epiplasm decorates the boundary of the cortical unit. The basal body structure is associated with a set of three rootlets. Two of rootlets are microtubule based rootlets. SR is the largest rootlet that originates from the basal body and extends several anterior basal body units. ICL is the innermost cytoskeleton contractile network. All these cytoskeleton structures along with basal body unit and its associated rootlets form the highly organized cortex of the *Paramecium* which is characterized by straight rows of basal body units from the posterior pole towards the anterior pole of the cell.

Image is used from Hufnagel, 1969 with permission.

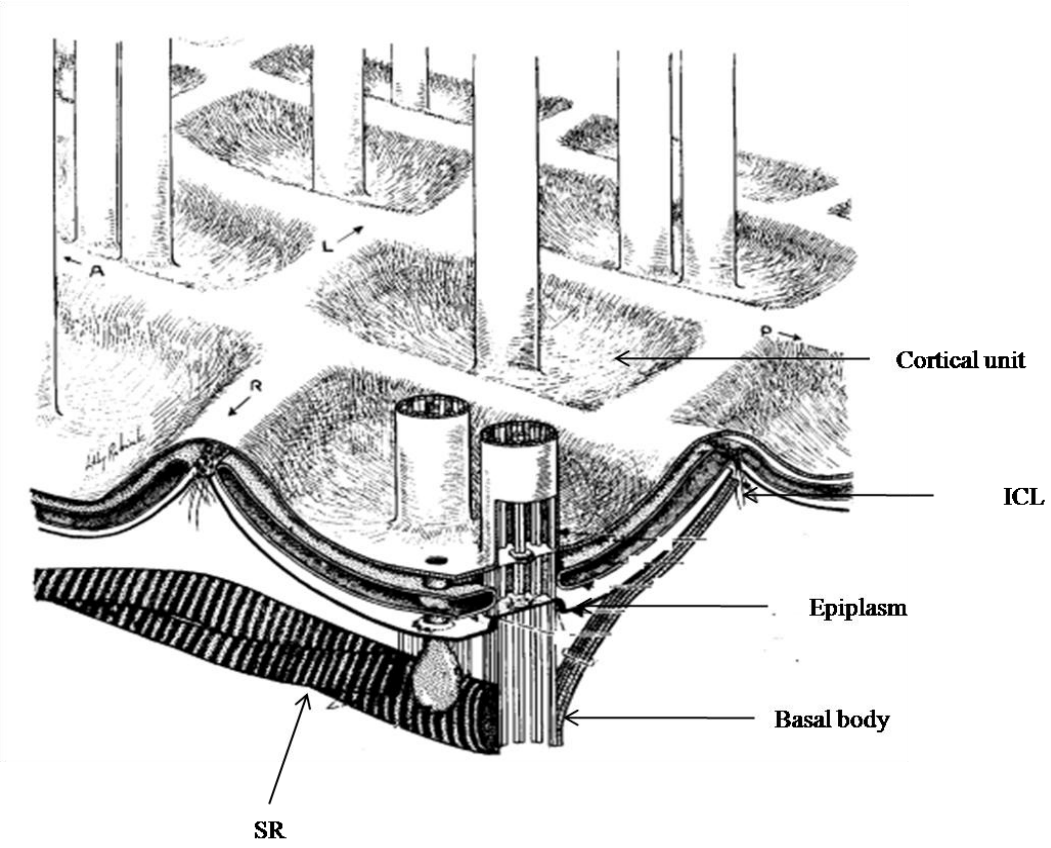

A
L
P
R
Cortical unit
ICL
Epiplasm
Basal body
SR

Fig. 1.7: The attachment points of SR, TR and PCR to the characteristic microtubule triplets of the basal body.

Rootlets attach to the basal bodies at the precise location to the microtubule triplets. The SR, PCR and TR rootlets attach to the microtubule triples 5, 9 and 4 of the basal body respectively. Image is used from Jerka-Dziadosz et al., 2013 with permission.

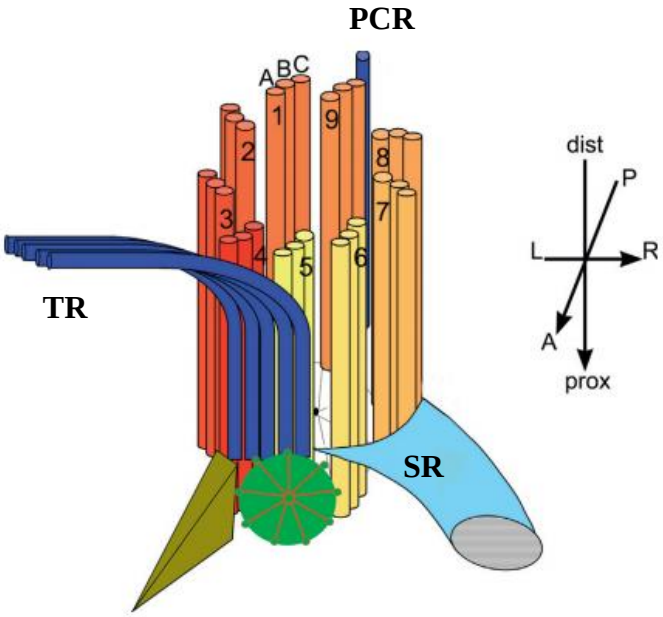
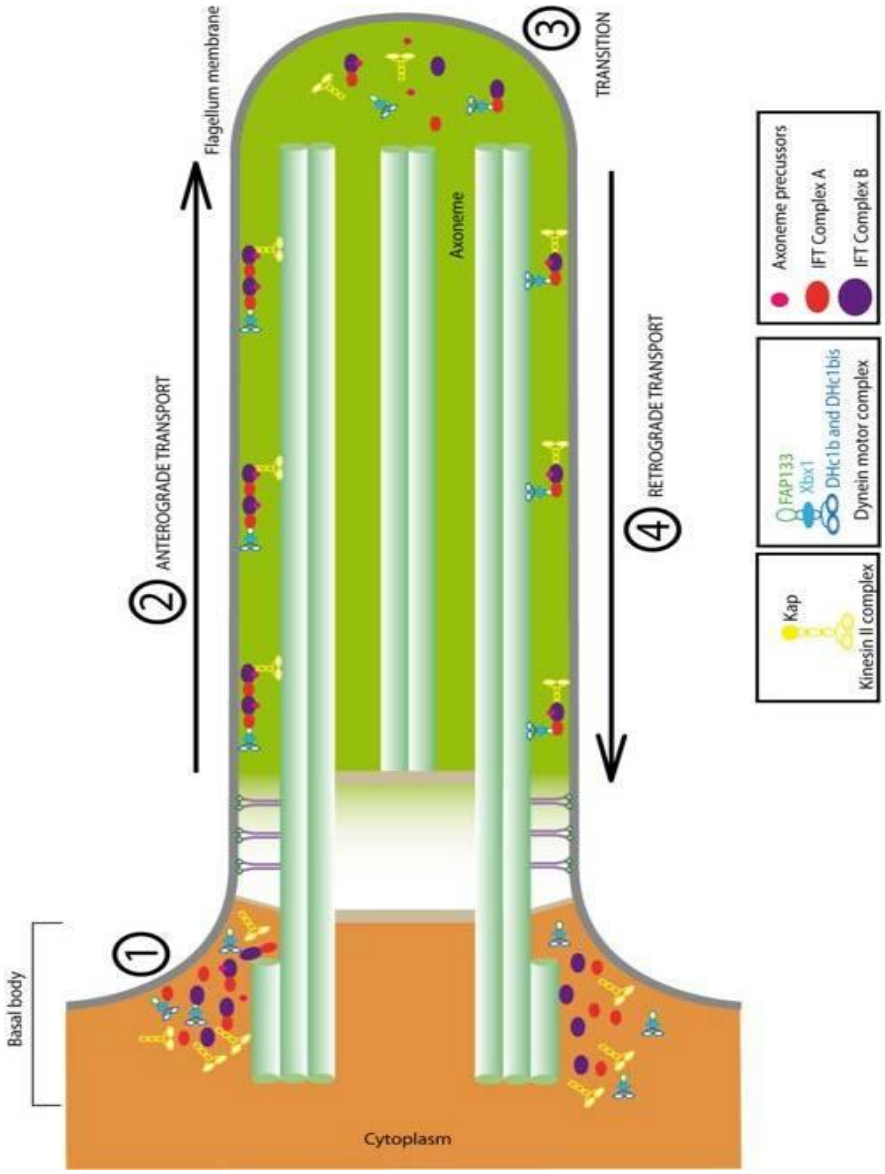


Fig. 1.8: An overview of intraflagellar transport (IFT).

An exclusive bi-directional transport system (IFT) is important for construction, maintenance and biogenesis of cilia. Anterograde movement of the IFT particles with cargo is facilitated by the kinesin motors from the base to the tip of the cilia. Retrograde movement of the IFT particles and cargo from the tip to the base of the cilia is powered by the dynein motors. Image used from Morga and Bastin, 2013 with permission.



Chapter 2: Multiple Functions of the Striated Rootlet (SR) Proteins of the *Paramecium* Basal Body

Introduction:

Cilia are hair-like projections extending from the surface of almost all human cell types. They perform a diverse set of functions including movement of mucous or spinal fluid, signaling needed for asymmetric development of organs, swimming of sperm, and signal transduction for normal development of bones, brain, kidneys (Fliegauf et al., 2007; Afzelius, 2004; Nakamura et al., 2006). An expanding group of autosomal or X-linked disorders collectively known as ciliopathies is caused by improper orientation or abnormal structure and function of cilia. A large number of symptoms, including cystic kidneys, central nervous system (CNS) malformations, polydactyly, respiratory illness, liver malformations, retinal degeneration, diabetes, obesity and skeletal dysplasias, characterize these disorders (Rayner et al., 1996; Afzelius, 2004; Badano et al., 2005; Badano et al., 2006). Failure of either nonmotile primary cilia or motile cilia to function results in severe disorders (Rayner et al., 1996; Badano et al.,2005; Bloodgood, 2010)

In this study we clarify the composition of the *Paramecium* ciliary basal body's striated rootlet (SR) and demonstrate that the SR plays a critical role in creating the orderly array of basal bodies in rows that run from pole to pole of the cell, likely through the interactions with centrins and other cytoskeletal elements underlying the cell surface.

Genes associated with the structure and functions of cilia are highly conserved, which makes it possible to use *Paramecium* with its thousand or more cilia per cell as a model to understand structure, function and development of cilia. An especially useful aspect of *Paramecium* as a model is its surface that is a pattern of repeating units, each

with one or two cilia arising (Vincensini et al., 2011). Deleterious changes that cause not just the loss of cilia but more subtle aspects like the anchoring or orientation of the cilia are visible upon by examination of the cell surface pattern of rows of basal bodies. Silencing human ciliopathy gene Meckelin (*MKS3*) in *Paramecium* caused the pattern of surface units and ciliary orientation to break down (Picariello et al., 2014). Rows of basal bodies became disoriented, surface units broke down and the SR of the basal body meandered under the surface. The motivation for this study was to understand the SR′s function in maintaining the organization of basal bodies and cortical units in rows and the relationship of the SR with the cortex cytoskeleton that underlies the cell surface and creates the cortical units.

The structure and the organization of the cilia have distinct features. The basal body with its nine microtubule triplets anchors the cilium at the surface. From the basal body arises the cilium with microtubule doublets of the axoneme covered with the ciliary membrane (Jana et al., 2014; Czarnecki and Shah, 2012; Kobayashi and Dynlacht, 2011).The *Paramecium* ciliary basal bodies align in straight rows of surface cortical units with one or two basal bodies per unit, coursing from the posterior to anterior pole of the cell. Each basal body connects to three rootlets on particular microtubule triplets that give the basal body an asymmetrical polarity. Two rootlets are microtubule based and project toward the posterior pole of the cell (Postciliary Rootlet - PCR) or laterally towards the adjacent rows of basal bodies (Transverse Rootlet -TR) (Allen, 1971; Iftode and Fleury-Aubusson, 2003; Jerka-Dziadosz et al., 2013).The third rootlet (SR) is a large structure of proteins unrelated to microtubules.

SRs are striated in ciliates (Rubin and Cunningham, 1973), amoeboflagellate (Larson and Dingle, 1981) and ciliated epithelial cells (Lemullois et al., 1991). The striated microtubule-associated fiber (System I fiber) of *Chlamydomonas* is the best studied SR. This SR comprises only a single polypeptide, SF-assemblin, a 34 kDa protein (284 amino acid residues) with an N-terminal short head domain (31 amino acid residues) and C-terminal helical rod domain, which has the ability to form a segmented coiled coil structure (Lechtreck and Melkonian, 1991; Weber et al., 1993). The presence of SF-assemblin in the SRs of green algae *Chlamydomonas* and *Spermatozopis similis* has been demonstrated by immunofluorescence and green fluorescent protein (GFP) epitope tagging (Lechtreck and Melkonian, 1991; (Lechtreck et al., 2002). The N-terminal head and C-terminal rod domains of the *Chlamydomonas* SF-assemblin protein play important roles in the formation of the striated microtubule associated fibers. Various deletions and insertions in both N-terminal head domain and C-terminal coiled coil rod domain change the solubility, striation pattern or overall appearance of the *Chlamydomonas* striated fiber (Lechtreck, 1998; Lechtreck et al 2002).

In *Paramecium*, the SR projects from the basal body toward the anterior of the cell past several more anterior basal bodies. The structure comprises proteins ranging 30-36 kDa molecular mass, some of which are phosphoproteins (Sperling, 1989; Sperling et al., 1991). These proteins interact to form a very long striated structure that is dynamic (i.e. changes length during the cell cycle) (Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003). The SR is assumed to play an important role in locking the basal body units in straight rows on the cell cortex through interaction with cytoskeleton structures.

These interactions could oppose the force generated by motile cilia that might move basal bodies and units out of alignment.

In this study, in order to understand the composition and function of the *Paramecium* SR, we identified the *SR* genes in the annotated genome and found 30 genes belonging to 13 Paralog Groups. Further analysis of the *SR* gene sequences for the potential ability of their protein products to form coiled coil domains combined with functional studies of the SR proteins using RNAi allowed us to classify these 30 proteins into five Structural Groups. All five Structural Groups contribute to the formation of SRs of appropriate length and striations. Loss of a single Structural Group causes the mis-orientation of basal bodies, mis-shapen cortical units, and misaligned SRs. Silencing of individual *SR* genes or SR Paralog Groups that are not also the Structural Groups does not have these effects.

Our N-FLAG epitope tagging data show that SR proteins distribute throughout the length of the SR. All the data presented in this study demonstrate an indispensable interaction between the SR and the innermost cytoskeleton structure, infraciliary lattice complex, of the cell cortex (Fig.2.2). We propose that the cortex centrin ICL1e subfamily proteins interact with the SR of the basal body unit to lock the basal body units into cortical rows at the cell cortex.

Materials and methods:

Stock, culture and chemicals:

Cells (stock 51s *P. tetraurelia*, sensitive to killer) were grown in wheat grass medium (Pines International, Lawrence, KS, USA) inoculated with *Aerobacter aerogenes* (Sasnar and Van Houten, 1989). All chemicals were purchased from Sigma-Aldrich (St Louis, MO, USA) unless otherwise noted.

SRs sequence analysis:

The SF-assemblin protein sequence from *Chlamydomonas reinhardtii* (Accession number: EDP05674.1) was used for searching the SR protein sequences in *Paramecium* annotated genome in the dedicated database ParameciumDB (http://paramecium.cgm.cnrs-gif.fr/). All possible SR protein sequences were checked in the NCBI conserved domain search and the Pfam database (http://pfam.xfam.org/) for the presence of conserved domains of SF-assemblin protein. Finally, the phylogenetic relationships among all the SR genes (nucleotide sequence) were analyzed using the MEGA6 software. (Tamura et al., 2013).

RNAi constructs:

SR RNAi constructs for 24 of 30 SR genes (*SR1a, SR1b, SR2, SR3, SR4, SR5a, SR5b,SR6a, SR6b, SR6d, SR7a, SR8a, SR8b, SR9, SR10a, SR10d, SR11a, SR11b, SR12a, SR12b, SR12c, SR13a, SR13c* and *SR13d*) and ICL RNAi constructs for three ICL genes (*Ptcen_icl1e, Ptcen_icl9a and Ptcen_icl10a*) (Genes ID number in ParameciumDB and genes nucleotide base positions in RNAi constructs are available in Table 2.1) were designed from the sequences in the *Paramecium* annotated genome using the

ParameciumDB database. In order to design specific constructs, we carried out off-target analysis for each construct as in ParameciumDB (http://paramecium.cgm.cnrsgif.fr/cgi/alignment/off- target.cgi). Specific oligonucleotide primers (Table 2.2) were used to amplify the designed sequence by using Genomic DNA as a template. The degree to which we were successful in targeting specific genes for silencing was verified through RT-PCR (see below).

Genomic DNA was purified by organic extraction as described in Picariello et al., 2014. PCR amplicons were then cloned directly into pCR2.1-TOPO vector (Invitrogen/Life Technologies), transformed into bacteria and sequenced according to the manufacturer (Invitrogen/Life Technologies) instructions. Correct insert sequences were cut from the pCR2.1-TOPO vector and ligated into the double T7-promoter vector L4440 (Addgene, Cambridge, MA, USA) using the Quick Ligation^{TM} reaction kit (New England BioLab Inc, Ipswich, MA, USA) as per the kit instructions. All vector inserts were sequenced. All SR RNAi plasmid constructs were maintained in *Escherichia coli* DH5α bacteria at -80^{0}C in a 30% glycerol stock.

RNAi feeding:

Escherichia coli strain HT115 (DE), which lacks RNaseIII, was used for RNAi feeding. HT 115 was transformed with 50ng of RNAi plasmid constructs of interest. HT115 bacteria transformed with L4440 with no insert were used as control. Overnight cultures of HT 115 bacteria transformed with a SR RNAi construct or ICL RNAi construct or control were used to inoculate 50 ml of Luria broth containing ampicillin (100 µg/ml) (LB-AMP) cultures. The bacterial cultures were incubated at 37^{0}C in a shaking incubator (New Brunswick Scientific) until they reached 595 nm optical

densities (OD) of 0.3-0.4. At the desirable OD, cultures were induced with isopropylthio-β-galactoside (IPTG) (RPI Corp., Mt. Prospect, IL, USA) by adding to a final concentration of 125 µg/ml. Cultures were then incubated at 37^0C in a shaking incubator for additional 4hrs to induce double stranded RNA synthesis. The induced bacteria cultures were centrifuged at 3,439xg for 10 minutes at 4^0C (Beckman Coulter, Brea, CA, USA) and the bacterial pellets were resuspended in 100 ml of wheat grass media containing additional stigmasterol (8µg/ml), ampicillin (100µg/ml) and IPTG (125 µg/ml). *Paramecium* cells which had recently undergone autogamy (checked by Dippell staining) (Dippell, 1955) were washed in Dryl's solution (1 mM Na_2HPO_4, 1 mM NaH_2PO_4, 1.5 mM $CaCl_2$, 2 mM Na-citrate; pH 6.8). For washing the paramecia, cells were harvested by centrifugation (500xg, Damon IEC Division Clinical Centrifuge, Needham Heights, MA, USA), resuspended in Dryl's solution and again harvested by centrifugation (500xg). Approximately 200-300 washed paramecia cells were added to the induced culture. All RNAi cultures were maintained at room temperature. An additional amount of stigmasterol (8µg/ml), ampicillin (100µg/ml) and IPTG (125 µg/ml) were added to the culture at 24hrs and 48hrs after feeding. Paramecia cells were collected by centrifugation (500xg) at 72hrs after feeding and subjected to further analysis unless noted otherwise. All the RNAi experiments were repeated a minimum of three times.

For isolation of the SRs, the RNAi treatments of cells were modified since large numbers of cells were required. The bacterial cells were induced in 500 ml of LB-AMP with IPTG (125 µg/ml) and the bacterial pellets resuspended in 1.5L of regular wheat grass medium containing stigmasterol (8µg/ml), ampicillin (100µg/ml) and IPTG (125 µg/ml). Approximately 6000 paramecia cells were added to the induced bacterial

cultures. Additional amounts of induced bacteria (500ml), ampicillin (100µg/ml) and IPTG (125 µg/ml) were added to the RNAi culture at 24, 48 and 72 hrs. Finally, paramecia cells were collected by centrifugation (500xg) at 96 hrs of RNAi feeding for the isolation of the SRs, immunofluorescence and the extraction of m-RNA.

Reverse transcription-PCR (RT-PCR)

We used RT-PCR to check the efficacy and specificity of the RNAi feeding procedure by following the protocol published by Yano et al., 2003. A pair of forward and reverse calmodulin primers (Table 2.2) was used in RT-PCR as control for the amount of input template, a check on the consistency of the method. Three different concentrations (undiluted, diluted 10 fold and 100 fold) of cDNA were used as template. All the experiments were repeated a minimum of three times. The results, while not quantitative, were highly reproducible. The Supplement images of the RT-PCR results are unaltered except cropped at the top and bottom the images. The images are each from a single experiment and controls are not shared among the images.

Immunofluorescence:

Paramecia cells after RNAi feeding (100 ml culture; Half of the culture was used for immunofluorescence, other half of the culture was used to extract mRNA for RT-PCR analysis) were harvested by centrifugation (500xg) (Damon IEC Division Clinical Centrifuge, Needham Heights, MA, USA). Cells harvested from 50 ml culture were resuspended in 100 ml of Dryl's solution and again harvested by centrifugation (500xg). This procedure was repeated three times in Dryl's solution. Finally, harvested cells were resuspended in ~100µl of Dryl's solution. Cells were permeabilized in PHEM buffer (60

mM piperazine ethanesulfonic acid (PIPES), 25 mM hydroxyethyl piperazineethanesulfonic acid (HEPES), 10 mM ethylene glycol tetraacetic acid (EGTA), 2 mM $MgCl_2$ and 0.1% Triton X-100; pH 6.9) and fixed for 30-40 min with fixation buffer (2% or 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA, USA), 2 mM $NaH_2PO_4•H_2O$, 8 mM Na_2HPO_4, 150 mM NaCl; pH 7.5). Cells were washed three times in blocking buffer (2 mM $NaH_2PO_4•H_2O$, 8 mM Na_2HPO_4, 150 mM NaCl, 10 mM EGTA, 2 mM $MgCl_2$, 0.1% Tween20, 1% BSA; pH 7.4) by resuspending the cells in blocking buffer followed by harvesting the cells by centrifugation (500xg). Immunostaining of the cells was done by incubating at room temperature for 1 hr in primary antibody. All antibodies were diluted in blocking buffer.

For the visualization of the SRs and basal bodies, primary antibodies were as follows: rabbit anti-SR at a dilution of 1:400 (gift from Janine Beisson, Centre de Génétique Moléculaire, Gif-sur-Yvette, France) and mouse anti-Glu-α-tubulin at a dilution of 1:500 (Synaptic Systems, Göttingen, Germany). For the visualization of cortical units with SRs, primary antibodies were as follows: mouse anti-2F12 at a dilution of 1:200 (gift from Jean Cohen, Gif-sur-Yvette, France) and rabbit anti-SR at a dilution of 1:400. For the visualization of epiplasm with basal bodies, primary antibodies were as follows: mouse anti-CTS32 (the antibody recognizes many epiplasmin proteins in *Paramecium*, gift from Anne Aubusson-Fleury, Gif-sur-Yvette, France) and rabbit anti-centrin at a dilution of 1:1000 (anti-Tetrahymena centrin, gift from Mark Winey, University of Colorado, Boulder, CO, USA). For the visualization of ICL with basal bodies and SRs, primary antibodies were as follows: mouse anti-FLAG M2 clone (since we used FLAG-Ptcen15 expressing cells for the RNAi experiments) at a dilution of 1:300

dilution (Sigma-Aldrich, St Louis, MO, USA), rabbit anti-centrin at a dilution of 1:1000 and rabbit anti-SR at a dilution of 1:400. Cells were washed three times in washing buffer (2 mM $NaH_2PO_4 \cdot H_2O$, 8 mMNa_2HPO_4, 150 mM NaCl, 0.1% Tween20; pH 7.4) by resuspending the cells in washing buffer followed by harvesting the cells by centrifugation (500xg) after primary antibody incubation. After washing, cells were incubated in secondary antibody at room temperature for 1hr. Secondary antibodies were Alexa Fluor 488 goat anti-mouse, Alexa Fluor 555 goat anti-mouse, Alexa Fluor 488 goat anti-rabbit, or Alexa Fluor 568 goat anti-rabbit (Molecular Probes/Invitrogen, Grand Island, NY, USA). After incubation in secondary antibodies, cells were washed three times with washing buffer (2 mM $NaH_2PO_4 \cdot H_2O$, 8 mM Na_2HPO_4, 150 mM NaCl, 0.1% Tween20; pH 7.4) as described above and suspended in Vectashield® mounting medium (Vector Labs, Burlingame, CA, USA) for imaging. All the images were captured using the DeltaVision® restoration microscopy system (Applied Precision), consisting of an inverted Olympus IX70 microscope (Olympus America, Center Valley, PA, USA) and a Kodak CH350E camera (Rochester, NY, USA). The Delta Vision® restoration microscopy system was supported by NIH Grant Numbers 5 P30 RR032135 from the COBRE Program of the National Center for Research Resources and 8 P30 GM103498 from the National Institute of General Medical Sciences.

For the visualization of the basal bodies along with all three rootlets, cells were deciliated as follows using trituration. The harvested cells were washed and re-suspended in 100 ml of Dryl's solution as described above. The Cells were collected by centrifugation in a clinical centrifuge (500xg) and resuspended in 9 ml of KCl buffer solution (5 mM KCl, 1 mM citric acid, 1 mM $Ca(OH)_2$, 1 mM Tris-base; pH 7.4). The

cell suspension was kept at room temperature for 5 minutes and centrifuged in a clinical centrifuge (500xg) to collect the cells. Two ml of KCl buffer was used to resuspend the harvested cells and absolute grade ethanol (≥99.5% (v/v)) was added rapidly to a final concentration of 5 % (v/v). A Pasteur pipet was used to triturate the cell suspension for 15 times immediately to sheer off the cilia. After trituration, the cell suspension was kept at room temperature for 5 minutes. The deciliated cells were collected from the bottom and resuspended in 100µl PHEM buffer for immunostaining as described above. Primary antibodies for the immunostaining of deciliated cells were as follows: rabbit anti-SR at a dilution of 1:400 (gift from Janine Beisson, Centre de Génétique Moléculaire, Gif-sur-Yvette, France), mouse anti-Glu-α-tubulin at a dilution of 1:500 (Synaptic Systems, Göttingen, Germany) and mouse anti-α-tubulin at a dilution of 1:200 (Sigma-Aldrich, St Louis, MO, USA).

Plasmid injection:

We used N-terminal FLAG pPXV plasmid (courtesy of Dr. W. John Haynes, University of Wisconsin, Madison, WI, USA) to clone one *ICL* gene and 13 different *SR* genes. All the gene sequences are available in ParameciumDB (http://paramecium.cgm.cnrs-gif.fr/). Accession numbers: *Ptcen15*:GSPATG0000107500; *SR1a*:GSPATG00025723001; *SR2*:GSPATG00022867001; *SR3*:GSPATG00014806001; *SR4*:GSPATG00008626001; *SR5a*:GSPATG00026454001; *SR6b*:GSPATG00014274001; *SR7a*:GSPATG00036966001; *SR8a*:GSPATT00032447001; *SR9*:GSPATT00011977001; *SR10a*:GSPATG00034929001; *SR11a*:GSPATG00002032001; *SR12a*:GSPATG00001334001; *SR13a*:GSPATG00025686001. All the primers used for

FLAG-epitope tagging are listed in Table 2.2. The target gene sequence was amplified by using Q5TM Hot Start high fidelity DNA polymerase (New England BioLab Inc, Ipswich, MA, USA) according to the manufacturer's protocol. Resulting amplicons were inserted into the pPXV plasmid using restriction enzymes (NheI/KpnI or ApaI/SacI) (New England BioLab Inc, Ipswich, MA, USA) and the amplicon sequences were confirmed by sequencing. All the FLAG pPXV plasmids containing the target gene sequences were linearized by using Not I restriction enzyme (New England BioLab Inc, Ipswich, MA, USA). The linearized plasmid was purified and resuspended in water at a concentration of 5-10 µg/µl. Approximately 10-15 wild type 51s *P. tetraurelia* cells which had recently undergone autogamy were injected with 5 to 50 pg of the linearized plasmid into the macronucleus. Individual cells were transferred to 500 µl of culture fluid in depression slides after injection and incubated at room temperature for 24hrs to recover and divide. Cells in each depression were then transferred to 10 ml culture fluid and maintained at 15^0C as individual clones. The presence of pPXV-3XFLAG-SR or pPXV-3X FLAG-Ptcen15 in individual clones was tested using PCR with extracted genomic DNA as a template.

Localization study:

For the localization study of SR protein, an individual clone expressing FLAG-SR was grown in 50 ml wheat grass culture media at 22^0C for 48hrs. Cells were immunostained and imaged as described above. Primary antibodies for the immunostaining were as follows: for FLAG-SR: mouse anti-FLAG M2 clone at a dilution of 1:300 (Sigma-Aldrich, St Louis, MO, USA) and for basal bodies: rabbit anti-centrin (anti-Tetrahymena centrin, gift from Mark Winey, University of Colorado,

Boulder, CO, USA) at a dilution of 1:1,000. All the images were captured using the same microscope system as described before. We followed the same procedure for 13 representative SR proteins to be checked for localization.

SR isolation:

For the isolation of SRs, we followed the protocol published by Sperling *et al.,* 1989 with modifications. Cells were harvested from 6L of culture fluid (4000-6000 cells/ml) by continuous flow centrifugation (IEC clinical centrifuge, 300xg) and washed three times in TEK buffer (20 mM Tris, 5 mM EGTA, 100 mM KCl; pH7) by resuspending the cells in 100 ml of TEK buffer followed by harvesting the cells by centrifugation (500xg). For the preparation of the crude cell cortex (without membrane), washed cells were resuspended in PHEM-1% Triton X100 in 10:1 volume ratio (buffer: cell pellet) and incubated on ice for 10 minutes. The protease inhibitors leupeptin (final concentration 5 µg/ml), pepstatin (final concentration 10 µg/ml) and PMSF (final concentration 1mM/ml) were added to all buffers. After incubation, cells were collected by centrifugation in a clinical centrifuge for 5min at 2000 rpm (650xg) and resuspended in 2 volumes of PHEM buffer without TritonX100. The cell suspension was then placed in a 55 ml glass-Teflon homogenizer (Wheaton, Millville, NJ, USA) and homogenized (~15-20 strokes) to make crude cortex. After homogenization, the crude cortex pellet was collected by centrifugation in a clinical centrifuge (1000xg) for 5 minutes. All subsequent steps were carried out at 4^0C. The crude cortex pellet was resuspended in 10 ml of PHEM buffer containing 100 mM $MgCl_2$ and homogenized very extensively (at least 70-80 strokes). After homogenization, the suspension was diluted in 1:2 volume ratio (homogenized suspension: PHEM buffer) using PHEM buffer and centrifuged at 4000

rpm (Beckman JA-20, 1940xg) for 7 minutes to sediment nuclei and trichocysts, which were then discarded. The collected supernatant was centrifuged at 10000 RPM for 15 minutes (in Beckman JA-20) to pellet the crude SR. The crude SR pellet was resuspended in 600 µl PHEM buffer and layered on PHEM-Optiprep density gradient. The PHEM-Optiprep gradient was made of 40%, 35%, 30%, 25%, 20%, 15% and 0% Optiprep (500 µl of each) in PHEM buffer and centrifuged in Beckman Coulter ultracentrifuge at 45000 rpm (SW 60 Ti rotor) for 2hrs. Each Optiprep-PHEM layer in the gradient was collected separately after centrifugation and diluted 10 fold with membrane buffer (10 mM Tris-Cl, 10mM Tris-base, 50 mM KCl, 5 mM $MgCl_2$, 1 mM EGTA, pH 7.4). To remove the Optiprep from the proteins, the suspension was centrifuged at 20000 rpm (in Beckman JA-20, 48750xg) for 30 minutes. The pellet was resuspended in 100 µl PHEM buffer. The 30% Optiprep-PHEM layer in the gradient had a whitish diffuse band (presence of SRs in this fraction was confirmed by Western blot) and used for further analysis.

Negative staining and Transmission Electron Microscopy (TEM):

Negative staining and TEM of the isolated striated rootlets were done at the Delaware Biotechnology Institute. In brief, Carbon coated 400 mesh copper grids were rendered hydrophilic with a Pelco easiGlow Glow Discharge Cleaning System. The grids were floated on drops of sample for several seconds, washed on four drops of water and then negative stained with 2% uranyl acetate (aq). After drying, samples were examined with a Zeiss Libra 120 transmission electron microscope operating at 120kV. Images were acquired with a Gatan Ultrascan 1000 CCD camera. The grids were then stored in a grid box for imaging.

Western blots:

The proteins in the SR preparation from FLAG-SR expressing cells were separated by gradient (5-18%) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970) and transferred to BioTraceTM nitrocellulose blotting membrane (Pall Life Sciences, Ann Arbor, MI, USA) by trans-blotting (Transphor electrophoresis, Hoefer, San Francisco, CA, USA). Blots were blocked in the buffer containing 0.5 g skim milk powder, 200 µl of Teleost fish gelatin, and 300 µl of normal goat serum (Vector Labs, Burlingame, CA, USA) dissolved in 10 ml of TBS-T (15 mM Tris,140 mM NaCl, 0.1% Tween, pH 7.5) at room temperature for 1 h with gentle rocking. Blots were probed using the rabbit anti-FLAG at a dilution of 1:2500 (Sigma-Aldrich, St. Louis, MO, USA) for overnight at 4^0C. Secondary antibody was alkaline phosphatase (AP) conjugated goat anti-rabbit (Sigma-Aldrich, St. Louis, MO, USA) at a dilution of 1:10000. Blots were developed with Nitro Blue Tetrazolium /5-Bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) (Moss Inc., Belfast, ME, USA).

Silver stained gels and mass spectrometry analysis:

The gradient (5-18%) SDS-PAGE gel after running the samples was silver-stained using directions of the FASTSilver^{TM} kit (G-Biosciences, St. Louis, MO, USA). We cut the gel pieces in the silver-stained gel, which represented the corresponding band in the Western blot for FLAG-SR protein in the 30% Optiprep fraction. We prepared the mass spectrometry samples as described before (Valentine et al., 2012). The prepared peptide samples were dissolved in 7 µl 0.1% formic acid and 2.5% acetonitrile, and 2 ul

were analyzed on the Thermo Q-Exactive mass spectrometer coupled to an EASY-nLC system (Thermo Fisher). Peptides were separated on a fused silica capillary (12 cm x 100 um I.D) packed with Halo C18 (2.7 um particle size, 90 nm pore size, Michrom Bioresources) at a flow rate of 300 nl/min. Peptides were introduced into the mass spectrometer via a nanospray ionization source at a spray voltage of 2.2 kV. Mass spectrometry data were acquired in a data-dependent top-10 mode, and the lock mass function was activated (m/z, 371.1012). Full scans were acquired from m/z 350 to 1,600 at 70,000 resolutions (automatic gain control [AGC] target, 1e6; maximum ion time [max IT], 100 ms; profile mode). Resolution for dd-MS2 spectra was set to 17,500 (AGC target: 1e5) with a maximum ion injection time of 50 ms. The normalized collision energy was 27 eV. A gradient of 0 to 40% acetonitrile (0.1% FA) over 55 min was applied. The spectra were searched against the Ptetraurelia_peptides_v1.99.13 protein database (http://paramecium.cgm.cnrs gif.fr/download/ fasta/) by Proteome Discoverer (PD) 1.4. The search parameters permitted a 10 ppm precursor MS tolerance and a 0.02 Da MS/MS tolerance. Carboxymethylation of cysteines was set up as fixed modifications and Oxidation of methionine (M) was allowed as variable modifications. Up to three missed tryptic cleavages of peptides were considered with the false-discovery rate set to 1% at the peptide level.

Swimming behavior analysis:

Swimming behavior was analyzed using dark field microscopy. Both the control cells or SR Paralog Group or SR Structural Group depleted cells were harvested by centrifugation after 72 hrs of RNAi construct feeding (500xg, Damon IEC Division Clinical Centrifuge, Needham Heights, MA, USA). Harvested cells were resuspended in

5mM KCl buffer (1 mM citric acid, 1 mM $Ca(OH)_2$, and 1 mM Tris-Base, pH 7.02 with Tris-Base). Swimming trajectories of the cells was captured using a dissecting microscope (Bausch & Lomb). In Brief, cells in 5 mM KCl buffer were placed in a small drop on a dissecting microscope. Another drop of 5 mM KCl buffer was placed near the drop containing the cells. A connection was made between the two drops using a Pasteur pipette. Finally, dark field images of tracks of swimming *Paramecium* cells were captured using a cannon DS digital camera with an exposure time of 2s.

Results

SR genes in *Paramecium*:

The SR of *Chlamydomonas* is the best studied of the ciliary SRs (Lechtreck and Melkonian, 1991). Using the single *Chlamydomonas SR* gene SF-assemblin to search in ParameciumDB, we identified thirty *Paramecium tetraurelia* genes in 13 Paralog Groups with one to eight genes in each. All thirty proteins translated from the nucleic acid sequences have the characteristic domains (N-terminal head domain and C-terminal rod domain) of *Chlamydomonas* SF-assemblin protein, confirmed by the NCBI conserved domain search and Pfam database. Evolutionary analyses of the thirty genes were conducted in MEGA6 (Tamura et al., 2013) (Fig.2.1A). The evolutionary history was inferred using the Neighbor-Joining method (Saitun and Nei, 1987). The MEGA 6 software uses the Maximum Composite Likelihood method to measure evolutionary distances. The evolutionary distances describe the units of base substitutions per site. The analysis involved all 30 nucleotide sequences. The final dataset includes 715 positions and eliminates all positions having the gap and missing data sets.

An analysis of the thirty *SR* genes' nucleic acid and amino acid sequences shows several layers of organization. We use the nomenclature *SR* for a gene name as distinguished from a Paralog Group or Structural Group in Fig.2.1. We use the term Paralog Group for duplicated genes because unlike ohnologs, they may not have exactly the same function. For example, *SR1a* is a gene and an ohnolog of *SR1b,* i.e. a duplicate

from whole genome duplication. Together they form Paralog Group SR 1 and also Structural Group 1. In another example, ohnologs *SR8a* and *SR8b* form Paralog Group SR 8; *SR10a, SR10b, SR10c* and *SR10d* are ohnologs that form Paralog Group SR 10. With *SR9* (Paralog Group SR 9), these two Paralog Groups form Structural Group 3.

The five Structural Groups are based on the hypothetical translation and location of heptamers that make up coiled coil domain (Fig.2.1B). Structural Groups 1 and 2 have two members whereas Structural Groups 3, 4 and 5 have seven, ten and nine gene members respectively. All the *SR* genes in the different Structural Groups code for proteins with two predicted coiled coil domains except the SR proteins in the Structural Group 5 that have only one coiled coil domain. These domains were identified by the program SMART (Letunic and Bork, 2017) and COILS (Lupas et al.1991). The position and the length of the coiled coil domains in the SR proteins within a particular Structural Group are similar but vary considerably between Structural Groups (Table 2.3). In addition, the amino acid sequences alignment of coiled coil domains of different Structural Groups clearly demonstrates amino acid sequences of the coiled coil domains of SR proteins within a Structural Group are very similar (supplement Fig.2.S1). But the amino acid sequences alignments of different SR proteins in different Structural Groups are not similar (supplement Fig.2.S2).

In ParameciumDB, some additional genes are described as potential SR candidates. We named them SR-like (SRL) since the encoded proteins do not have the characteristic domains of *Chlamydomonas* SF-assemblin protein. We used MEGA6 software to make a phylogenetic tree, which includes both the SRs and SRLs nucleotide sequences (supplement Fig.2.S3).The nucleotide sequences of the SRL genes (sequences

that are available in ParameciumDB and used for phylogenetic analysis) are very different from the SR genes based on the size and nucleotide sequence similarity (supplement Fig.2.S3).

Localization of Flag-SR proteins:

From the thirty *SR* genes in thirteen Paralog Groups, we selected one randomly from each of the Paralog Groups (*SR1a, SR2, SR3, SR4, SR5a, SR6b, SR7a, SR8a, SR9, SR10c, SR11a, SR12d* and *SR13c*) for FLAG epitope tagging. Wild type *P. tetraurelia* cells injected with the 5′-3xFLAG plasmid to express FLAG peptide served as control cells. *Paramecium* cells injected with the representative 5′-3xFLAG-*SR* plasmids to express representative FLAG-SR peptides were used to observe the location of different SR proteins. All the cells were treated with anti-FLAG (expressed FLAG or FLAG epitope tagged SR proteins) and anti-centrin (basal bodies) except in the uninjected wild type *Paramecium* control cells (Fig.2.3). Wild type cells were treated with anti-SR and anti-1D5 to visualize the SRs (red) and basal bodies (green) respectively. Staining pattern of SRs along with basal bodies in wild type cells allowed us to demonstrate the location of different FLAG-SR proteins in FLAG-SR expressing cells.

In wild type cells, basal bodies (green in wild type control cell) are organized in straight rows between the posterior pole and the anterior pole of the cell; SRs (red in wild type control cell) emanate from the basal body, extend towards the next anterior basal body of the rows and traverse several basal body units toward the anterior (Fig.2.3). The basal body staining pattern is the same in the FLAG expressing control cells or the control non-expressing wild type cells, demonstrating that FLAG epitope tag expression alone does not affect the highly organized basal body rows in *Paramecium*. In addition,

in the FLAG expressing control cells, there are almost no signals for the FLAG, making it possible to interpret the immunofluorescence from FLAG-SR expressing cells (Fig.2.3). FLAG staining in the cells expressing FLAG-SR proteins (SR1a, SR2, SR3, SR4, SR5a, SR6b, SR7a, SR8a, SR9, SR10c, SR11a, SR12d or SR13c) shows that FLAG-SR proteins are throughout the SRs (Fig.2.3). It appears, within the limits of the fluorescence microscopy (resolution 200 nm), that the proteins are distributed from base to tip. In addition, SRs (green in FLAG-SR expressing cells) in all the representative FLAG-SR expressing cells are similar in appearance to the SRs (red) in the control non-expressing wild type cells.

Proteins without the characteristic domains of the SF-assemblin do not localize in the SRs (supplement Fig. 2.S4). We also checked the location of these proteins since the ParameciumDB describes them as potential candidates of the SRs. Immunostaining of FLAG-SRL proteins expressing cells with anti-FLAG describes FLAG-SRLs are not found in the same locations as FLAG-SR proteins. FLAG-SRL1a and FLAG-SRL3a locate intracellularly, in the basal bodies, and throughout the cilia respectively. Both FLAG-SRL4a and FLAG-SRL5a locate in the epiplasm (supplement Fig.2.S4).

Depletion of SR Structural Group leads to basal body row misalignment and abnormal SR appearance:

We depleted mRNAs for *SR* genes by RNAi and checked the phenotypes for basal body rows and SRs alignment. The RNAi construct designs were very crucial. Since nine Paralog Groups among thirteen have more than one gene ((four Paralog Groups (SR 2, SR 3, SR 4 and SR 9) have only one gene)), RNAi experiments were designed to deplete the messages of all the genes within a targeted Paralog Group but not affect the message

of other nontargeted Paralog Groups. We confirmed the level of message depletion within a targeted Paralog Group without affecting the nontargeted Paralog Groups by RT-PCR. Analysis of RT-PCR data from various RNAi experiments clearly demonstrates our designed RNAi constructs only depleted the message for the targeted Paralog Group without affecting the message of the nontargeted Paralog Groups (supplement Fig.2.S5). All experiments with their accompanying RT-PCR were repeated three times.

The phenotypes of the L4440 fed control cells or the RNAi cells were examined with immunofluorescence. In all cells, anti-1D5 marks the basal bodies and anti-SR marks the SRs. The whole cell images (Fig.2.4) comprise the stack of Z-sections approximately 10 μm thick to ensure all basal bodies along with SRs are visible. In L4440 fed control cells (Fig.2.4A), basal bodies (green) and SRs (red) have a highly organized characteristic pattern. Basal bodies are in straight rows from the posterior pole to the anterior pole both on the dorsal and the ventral surfaces of the cell. SRs originate from the basal bodies and extend toward the anterior pole of the cell and traverse several basal body units (Fig.2.4B).

We took advantage of genes in which the Paralog Group and Structural Groups were the same. The ohnologs *SR7a* and *SR7b* are so close in sequence that we could silence them together. Together they are the only members of both the Paralog Group SR 7 and Structural Group 2. RNAi for *SR7a/b* depletes mRNA for the entire Structural Group 2, and severely disrupts the basal body rows alignment; SRs orientation and appearance (Fig.2.4C, 2.4D). Misaligned basal bodies and abnormal SRs are found all

over the cell surface except the oral groove area, suggesting that both the dorsal and ventral surfaces of the cells are similarly affected.

In another pair of Paralog Group genes, the ohnologs are *SR1a* and *SR1b*. Both of the genes are the member of Paralog Group SR 1 and Structural Group 1. The *SR1a* and *SR1b* genes are not too close in sequence to deplete with RNAi of either *SR1a* or *SR1b* alone. When they are treated by RNAi together (RNAi constructs for both *SR1a* and *SR1b*), the entire Paralog Group SR 1 or Structural Group 1 is reduced and the dramatic phenotype of misaligned basal bodies and abnormally oriented and shaped SRs is seen (Fig.2.5L).

The dramatic phenotype of Structural Group 1 (Paralog Group: SR 1) or Structural Group 2 (Paralog Group: SR 7) depleted cells does not appear upon the depletion of other SR Paralog Group unless the entire Structural Group gene transcripts are depleted together. For example depletion of genes that are the only members of their Paralog Groups but not the only members of their Structural Groups produces no obvious phenotype: SR 2, SR 3, SR 4 and SR 9 (Fig.2.5A, 2.5B, 2.5C and 2.5D).

Depletion of entire Paralog Groups also does not create the dramatic phenotype of basal body rows misalignments if the Paralog Group is not the only Paralog Group in the Structural Group. RNAi for the following depleted entire Paralog Groups as shown by RT-PCR but result in no phenotype: SR 5 (ohnologs *SR5a,b*) (Fig.2.5E), SR 6 (ohnologs *SR6a,b,c,d*) (Fig.2.5F), SR 8 (ohnologs *SR8a,b*) (Fig.2.5G), SR 10 (ohnologs *SR10a,b,c,d*) (Fig.2.5H), SR 11 (ohnologs *SR11a,b)* (Fig.2.5I), SR 12 (ohnologs *SR12a,b,c,d*) *(*Fig.2.5J), or SR13(ohnologs *SR 13a,b,c,d*) (Fig.2.5K).

These results are in extreme contrast with the results when entire Structural Groups are depleted: Structural Group 5 (Paralog Groups: SR 2, SR 3, SR 4, SR 5 and SR 6) (Fig.2.5M), Structural Group 3 (Paralog Groups: SR 8, SR 9 and SR 10) (Fig.2.5N) or Structural Group 4 (Paralog Groups: SR 11, SR 12 and SR 13) (Fig.2.5O). These are in addition to the previous Structural Group 1 (Paralog Group: SR 1) and Structural Group 2 (Paralog Group: SR 7) that are shown in Fig.2.4C; 2.4D and Fig. 2.5L. The whole cell images are available in the supplement Fig.2.S6.

Isolation of SRs and Mass Spectrometry analysis:

We used Optiprep density gradients to concentrate a cellular fraction from the *Paramecium* cell cortex that included SRs (Fig. 2.6A). Fig. 2.6B shows the SDS-PAGE analysis and silver stained gel of all of the Optiprep fractions. Clearly the proteins that can be identified by silver staining are almost exclusively found in the 30% fraction. What is left of the source of the proteins, the cell cortices, is pelleted at the bottom of the centrifuge tube and is not loaded on the gel. To examine further whether the SRs are in the 30% faction, we performed this preparation on cells expressing FLAG-SR8a. An immunoblot with anti-FLAG antibody clearly shows that there is a band of the expected size (~33 kDa) in the 30% Optiprep fraction (Fig.2.6C). To take this analysis further, we embedded the proteins from the 30% fraction from FLAG-SR8a expressing cells in low melting agarose and used anti-FLAG antibody to visualize the structure (Fig.2.6D). This structure has the size (~5 µm) and shape of an SR. We obtained the same result for microscopic analysis, SDS-PAGE followed by silver staining and immunoblotting with anti-FLAG antibody using FLAG-SR7a or FLAG-SR2 expressing cells.

Using the band on the immunoblot as a guide to the location of SRs in the silver stained gels and the 30% Optiprep fraction, we excised the portion in the silver stained gel pieces corresponding to the location of the band (indicated by the red box in Fig.2.5B). LC-MS/MS analysis confirms the presence of SR proteins (Table.2.4) in the excised area of the gel. In Table 2.4, 'Total Peptide Sequence Identified' shows the total number of the peptide sequences identified in LC-MS/MS analysis for a particular Structural Group and 'Unique to the Paralog Group' indicates the number of peptide sequences identified uniquely for a specific member of the Paralog Group.

For Structural Group 1, eight and thirteen unique peptide sequences identified for the SR1a and SR1b respectively. Some peptides in common to both SR1a and SR1b were also identified. For Structural Group 2, only peptide sequences in common to SR7a or SR7b were identified, since both of the proteins have identical amino acid sequences. Structural Group 3, 4 and 5 showed a similar result for their group members. In these groups, mass spectrometry identified unique peptide sequences for some members and also those shared among all members. Unique peptide sequences were found for all the SR proteins in Structural Group 3, 4 and 5 except SR6a, SR6b, SR8b, SR12a, SR12b, SR12c, SR12d and SR13d.

In addition, we found additional proteins when we analyzed the areas of the gel outside the red box in Fig. 2.5B. Centrin family proteins and centrin binding proteins (which form the ICL network), are the most interesting proteins that co-isolated with SRs in this 30% fraction (supplement Table 2.S13). α-tubulin and β-tubulin proteins subfamilies are also found and probably signify that occasionally the basal bodies remain

associated with the SRs. Heat shock proteins and ribosomal proteins are likely to be contaminants of the preparation.

In other experiments, we depleted the Structural Group1 (Paralog Group: SR 1), isolated the SRs and analyzed by LC-MS/MS. A detailed list of proteins identified by mass spectrometry is described in supplement Table 2.S14. As expected (since we depleted the transcripts of Structural Group 1: Paralog Group SR 1) mass spectrometry failed to identify any peptide sequence specific for the Structural Group 1 in the SRs preparation isolated from SR Structural Group 1 depleted cells. Interestingly, centrin family proteins and centrin binding proteins (which form the ICL network) are also not identified in these experiments. We observed similar results (absence of SR Structural Group 2 proteins, centrins and centrin binding proteins) in the LC-MS/MS analysis of the isolated SRs from SR Structural Group 2 depleted cells. The list of proteins identified by mass spectrometry in the SR preparation isolated from SR Structural Group 2 depleted cells is available in supplement Table 2.S15.

Abnormal SRs appearance correlates with the basal bodies misalignment in SR Structural Group depleted cells:

We used negative staining and transmission electron microscopy to examine SRs in the 30% Optiprep fraction from control, SR Paralog Group or SR Structural Group depleted cells. In control cells, the SR is a long (average length of the SR is 5.5µm, Fig.2.7B), slightly bent fiber in which tapering starts at approximately 700nm (Fig.2.7A). Although length and thickness of the isolated SRs are variable, none of the SRs length is less than 3.5µm.

In the SR Paralog Group depleted cells (SR 2; SR 3; SR 4; SR 5; SR 6; SR 8; SR 9; SR 10; SR 11; SR 12or SR 13) in which basal bodies rows are in straight alignment, isolated SRs length is close to control cells (Fig.2.7A). Strikingly, depletion of SR Structural Group 1 (Paralog Group: SR 1), Structural Group 2 (Paralog Group: SR 7), Structural Group 3 (Paralog Group: SR 8-SR 10), group 4 (Paralog Group: SR 11-SR 13) or Structural Group 5 (Paralog Group: SR 2-SR 6) which cause misaligned basal bodies rows, also affect the length of the SRs. In these cells, isolated SRs are significantly shorter than control cells (Fig.2.7B).

We used transmission electron microscopy to examine the striation patterns of SRs. In control cells or SR Paralog Group depleted cells that retained the straight alignment of basal body rows, the striation pattern of their SRs showed typical periodicity in which major striation occurs at 24-36nm intervals (Fig.2.8A, indicated by red bracket). In cells depleted of the SR Structural Groups (Structural Group 1, 2, 3, 4 or 5), which all show misaligned basal body rows and abnormal SRs structure in immunostaining, the striation patterns have changed significantly (Fig.2.8B). Major striation periodicity of the isolated SRs in these cells is not quantifiable since in these isolated fibers identification of the major striation is difficult. Moreover, approximately forty to sixty percent of the isolated SRs in these affected cells have no striations at all (Fig.2.8B).

Altered swimming behavior is evident in the SR Structural Group depleted cells with misaligned basal bodies rows:

Given the severe distortion of basal body rows and SR alignment in the cells depleted of the SR transcripts by Structural Group (Fig.2.4 and 2.5), we asked whether the cells can swim and if so is it in a normal path. First, we confirmed with anti-α tubulin

staining of the cells that they are ciliated (not shown).To analyze the swimming behavior, we recorded swimming paths in 5mM KCl buffer solution using dark field microscopy. Each white line in the images is the 2 second swimming path of one cell (Fig.2.9). In 5mM KCl buffer solution, control cells swim in a straight line. Depletion of SR Paralog Groups (SR 2; SR 3; SR 4; SR 5; SR 6; SR 8; SR 9; SR 10; SR 11; SR 12 or SR 13) which do not cause basal body rows misalignment, swim in a straight path as do control cells (Fig.2.9A- L) . Interestingly, depletion of any SR Structural Group ((Structural Group1 (Paralog Group: SR 1), Structural Group 2 (Paralog Group: SR 7), Structural Group 3 (Paralog Group: SR 8-SR 10), Structural Group 4 (Paralog Group: SR 11-SR 13) or Structural Group 5 (Paralog Group: SR 2-SR 6)) which cause basal body rows misalignment (Fig.2.9M-Q), show very distinct swimming behavior compared to control cells.

Rootlet angles at the basal body remain constant in SR Paralog Group or SR Structural Group depleted cells:

The three rootlets of the basal body are at fixed angles. The SR emanates from the left side of the basal body and extends toward the anterior past several more anterior basal bodies. The PR that emanates from the basal body at 7 o'clock orientation with the anterior pole of the cell pointing 12 o'clock orientation extends toward the posterior pole of the cell. The TR that emanates from the basal body at 5 o'clock orientation with the anterior pole of the cell pointing 12 o'clock orientation extends perpendicularly towards adjacent basal bodies row. In a two basal body unit, both SR and PR associate with only the posterior basal body while TR associates with both the anterior and posterior basal bodies.

To analyze whether the misalignment of basal body rows and disordered organization of SRs were accompanied by changed angles (implying that their sites of attachment to the basal body were different from controls), we treated the cells with anti-α tubulin for the microtubule based rootlets (TRs and PRs), anti-1D5 for basal bodies and anti-SR for the SRs. The tubulin and the basal bodies had to be visualized with the same fluorophore on the secondary antibodies. Therefore, we deciliated cells just before immunostaining since one of the antibodies (anti-α tubulin) which stains the microtubule based rootlets also stains the cilia. Fields of basal bodies on the cell surfaces of different cells are shown in (Fig.2.10).

We examined angles for control, SR Paralog Group (SR 2) and SR Structural Groups (Structural Group 1: SR1 and Structural Group 2: SR 7) depleted cells. To analyze the angles between rootlets, we marked the basal bodies in Fig.2.10 with white circles having all three rootlets visible. We selected basal bodies in the affected surfaces area of the cells for SR Structural Group depleted cells. We then measured the angles between rootlets with a protractor (Fig.2.11A and 2.11B). The angles between PRs and the TRs remain consistent with the control, SR Paralog Group (SR 2) or SR Structural Group (Structural Group 1 or 2) depleted cells (Fig.2.11C). The angles between PRs and the SRs also remain constant in the control, SR Paralog Group (SR 2) or SR Structural Group (Structural Group 1 or 2) depleted cells (Fig.2.11D). We counted 100 basal bodies for each of the cell types.

Depletion of SR Structural Group causes distorted cortical units with abnormal SRs:

Immunostaining of cells with anti-2F12 and anti-SR antibodies decorates the ridges of the cortical units and SRs respectively. In control L4440 fed cells, cortical units along with SRs are organized in a highly ordered pattern on the surface of the cell (Fig.2.12A). All the cortical units are hexagon shaped bounded by ridges elevated above the center of the hexagon from which one or two cilia arise.

The cortical units are organized in straight rows from the posterior pole towards the anterior pole of the cell. Each SR originates from the basal body within a cortical unit, extends towards the anterior pole of the cell and transverses more than one anterior cortical unit (Fig.2.12B) (If there are two basal bodies in the unit, only the anterior basal body has the SR). The Paralog Group SR 2 depleted cells show the same pattern of the organization of the cortical units along with the SRs compared to control cells (Fig.2.12C, 2.12D). We observed same phenotype with depletion of transcripts from other *SR* genes or Paralog Groups (SR 2; SR 3; SR 4; SR 5; SR 6; SR 8; SR 9; SR 10; SR 11; SR 12 or SR 13) (supplement Fig.2.S7).

Cells depleted for SR Structural Group 2 (Paralog Group: SR 7) show multiple areas of distorted cortical units with a loss of alignment of the cortical units from the posterior pole towards the anterior pole of the cell (Fig.2.12E). SRs in these cells (in the area of distorted cortical units) are abnormal in appearance and they are mis-directed away from the axis of the posterior-anterior poles of the cell (Fig.2.12F). Abnormal appearance of SRs with distorted cortical units could be observed on the dorsal surface and ventral surfaces of the cell. We observed the same dramatic phenotype in all other SR Structural Group depleted cells ((Structural Group1 (Paralog Group: SR 1), Structural

Group 3 (Paralog Group: SR 8-SR 10), Structural Group 4 (Paralog Group: SR 11-SR 13) or Structural Group 5 (Paralog Group: SR 2-SR 6)) (supplement Fig.2.S7).

We used a second approach to visualize the cortical units for the effect of SR depletion by using antibodies to epiplasm, which is a layer immediately beneath the surface membrane of the *Paramecium* cortex (we used mouse anti-CTS32 which recognizes many proteins of the epiplasm, gift from Anne Aubusson-Fleury, Gif-sur-Yvette, France). Epiplasm appears as a territory within each cortical unit, with an opening through which the top of the basal body/cilium emerges (Allen, 1971; Nahon *et al*., 1993; Aubusson-Fleury *et al*., 2013). While the epiplasm does not appear in these ultra structure studies to interact directly with the SR or basal body, it is a part of the cortex and a marker of the repeating cortical unit structure. We had already observed that depletion of SR Structural Groups disrupts the ridges and shapes of the cortical units, we expected to also see disruption of the epiplasm territories that underlay the cortical units. Therefore, we used anti-epiplasmin and anti-centrin to visualize epiplasm territories and basal bodies respectively in control cells and those depleted of SRs.

In the control L4440 fed cells anti-epiplasmin antibody decorates the epiplasm, which segregates into regular shaped territories across the cell surface (Fig.2.13A, 2.13B). Each regular territory of epiplasm harbors one or two basal body unit at the center of the territory. Straight rows of basal bodies along with epiplasm are organized in the highly ordered pattern on the surface of the cell (Fig.2.13A, 2.13B). In the Paralog Group SR 2 depleted cells, epiplasm territories along with basal body rows appeared normal compared to the control cells (Fig.2.13C, 2.13D). We observed the characteristic pattern of straight basal body rows along with regularly shaped epiplasm territories in

cells depleted of all other SR Paralog Groups (SR 2; SR 3; SR 4; SR 5; SR 6; SR 8; SR 9; SR 10; SR 11; SR 12 or SR 13) (supplement Fig.2.S8).

In cells depleted of SR Structural Group 2 (Paralog Group: SR 7), the phenotype is very different. Basal body rows are misaligned along with very unusual pattern and shape of territories of epiplasm on both dorsal side and ventral side of the cell. These territories are highly irregular in shape in the areas of the cell surface that shows basal body row misalignment (Fig.2.13E, 2.13F). We observed the same irregular pattern of epiplasm territories along with misaligned basal body rows in all other SR Structural Group depleted cells ((Structural Group1 (Paralog Group: SR 1), Structural Group 3 (Paralog Group: SR 8-SR 10), group 4 (Paralog Group: SR 11-SR 13) or Structural Group 5 (Paralog Group: SR 2-SR 6)) (supplement Fig.2.S8).

Depletion of SR Structural Group affects the infraciliary lattice complex (ICL) organization:

The previous section describes our study of the effects on SR depletion on the cortical units by studying the ridges and epiplasm that define the shape of the cortical units. Here we examine what effect that SR depletion might have on a deeper layer of the cortex. The Infraciliary Lattice Complex (ICL) is the innermost cytoskeleton fibrous network of the *Paramecium* cortex (Allen 1971; Garreau de Loubresse et al., 1988; Garreau de Loubresse et al 1991; Gogendeau et al., 2008). ICL forms a contractile network and organizes into a bundle of filaments that does not exactly follow the boundary of the highly organized cortical units of the *Paramecium* cell surface (Allen 1971; Garreau de Loubresse et al., 1988; Garreau de Loubresse et al., 1991). The ultra structure study (Garreau de Loubresse et al., 1988; Garreau de Loubresse et al., 1991) of

the whole cell and partially purified ICL network demonstrates ICL network appears to have close association with basal body units and other cytoskeleton structure of the cell cortex.

To investigate how the depletion of SR proteins affects the ICL, we used RNAi with *Paramecium* cells expressing FLAG-Ptcen15 to visualize the ICL in the absence of an antibody for the ICL. (Ptcen15 protein is a centrin identified in the ICL network and a member of the ICL1e subfamily; Gogendeau *et al*., 2008). FLAG-Ptcen15 injected cells fed with L4440 construct were used as the control. We also used 5′-3xFLAG plasmid vector injected cells to feed L4440 plasmid construct or *SR* RNAi constructs which allowed us to demonstrate 5′-3xFLAG plasmid expression had no effect on the highly organized basal body rows or the ordered organization of the SRs. All the cells were treated with anti-FLAG, anti-centrin and anti-SR antibody to visualize the ICL complex, basal body units and SRs respectively.

In the control L4440 fed cells, the ICL network forms a polygonal mesh of fibrous network on the dorsal surface of the cell (Fig.2.14A). Basal body row alignment remains straight from the posterior pole towards the anterior pole of the cell. SRs emanate from the basal body unit, extend towards the anterior pole of the cell and have a highly ordered pattern (Fig.2.14B). The basal body unit and the proximal end of the SR in red demonstrate a close association with ICL complex (white arrow in Fig 2.14B). Overall, the SR structure maintains a close association with ICL mesh network throughout its entire length.

In Paralog Group SR 2 depleted cells, the ICL network meshes appear the same as in the control cell. Basal body rows along with the highly ordered SRs appear normal as

compared to control cells (Fig.2.14C).The basal body unit and SR maintain the close association with ICL network as evident in the enlarged image of the highlighted area of the cell (Fig.2.14D). We observed the same phenotype in the Paralog Group SR 10 depleted cells (supplement Fig.2.S9)

To analyze the ICL meshes network in SR Structural Group depleted cells, we depleted the SR Structural Group 2 (Paralog Group: SR 7) in FLAG-Ptcen15 expressing *Paramecium* cells. Immunostaining of the cell shows that the ICL network meshes somehow appear intact in the SR Structural Group 2 depleted cells. Basal body rows misalign from their straight rows; SRs have a disordered organization with abnormal appearance (Fig.2.14E, 2.14F). In the area of misaligned basal body rows with abnormal SRs, the honeycomb like polygonal territory of the ICL mesh network appears much more different or irregular as compared to the control cell (white arrow in Fig.2.14F). We observed the same phenotype in SR Structural Group1 (Paralog Group: SR 1) depleted cells (supplement Fig.2.S9)

In L4440 fed 5′-3xFLAG expressing cells, anti-FLAG staining show no green fluorescence from the ICL (Fig.2.14 G-J). Straight basal body rows along with the highly ordered organization of SRs from the posterior pole towards the anterior pole of the cell demonstrate that FLAG epitope tag expression does not affect the surface organization of the *Paramecium* (Fig.2.14G, 2.14H). Depletion of SR Structural Group 2 in 5′-3xFLAG expressing cells clearly demonstrates misaligned basal body rows with abnormal SRs appear due to depletion of a SR Structural Group (Fig.2.14I, 2.14J).

In a reciprocal depletion study, ICL 1e subfamily depleted cells show a dramatic phenotype of misaligned basal body rows with the disordered SRs:

We found in our analysis of SRs by mass spectrometry that ICL subfamily ICL1e proteins and centrin binding proteins co-isolate with SRs suggesting that they might be the ICL proteins that allow the SRs to anchor to the ICL. We have shown that the depletion of the SR Structural Groups cause distortion of the ICL network. In order to determine whether there is a reciprocal interaction of ICL network and SRs, we depleted the transcripts for the entire ICL1e subfamily and examined the effect on the ordered organization of the SRs. ICL network in *Paramecium* consists of ten subfamilies (ICL1a, ICL1e, ICL3a, ICL3b, ICL5, ICL7, ICL8, ICL9, ICL10 and ICL11) (Gogendeau *et al*., 2008). The rationale behind selecting ICL1e subfamily for transcripts depletion is that we found ICL1e subfamily proteins with all of the SR isolation preparations. We also depleted ICL9 and ICL10 subfamily transcripts as controls because these proteins were not identified through proteomics of the SR preparations.

The ICL 1e sub family in *Paramecium* includes seven genes such as *Ptcen_icl1e, Ptcen_icl1g, Ptcen8, Ptcen10, Ptcen12, Ptcen15* and *Ptcen18* (Gogendeau *et al*., 2008). The RNAi construct design to deplete the ICL1e subfamily was crucial. We used the *Ptcen_icl1e* gene sequence to create the RNAi construct. The RNAi experiment was designed to deplete the messages of all the seven genes within the targeted ICL 1e subfamily but not affect the message of another non-targeted subfamily such as ICL9 or ICL10. We confirmed the level of message depletion within the targeted ICL1e subfamily and non-targeted ICL9 and ICL10 subfamilies by RT-PCR. Analysis of RT-PCR data from various RNAi experiments clearly demonstrates our designed RNAi constructs only depleted the message for the targeted ICL1e subfamily without affecting the message of the non-targeted ICL9 and ICL10 subfamilies (supplement Fig.2.S10)

ICL9 and ICL10 subfamily includes four (*Ptcen_icl9a, Ptcen_icl9b, Ptcen_icl9c* and *Ptcen_icl9d)* and two (*Ptcen_icl10a* and *Ptcen_icl10b)* genes. We used *Ptcen_icl9a* and *Ptcen_icl10a* gene sequences to create the RNAi constructs to deplete the ICL9 and ICL10 subfamilies respectively. RT-PCR analysis clearly demonstrates our RNAi constructs depleted the transcripts for targeted subfamily without affecting the transcripts of the non-targeted subfamilies (supplement Fig.2.S11 and Fig.2.S12).

All the cells (L4440 fed control and RNAi cells) were treated with anti-1D5 and anti-SR to visualize the basal bodies and SRs respectively. ICL 9 (Fig.2.15A) or ICL 10 (Fig.2.15C,) subfamily depleted cells treated with anti-1D5 show straight basal body rows similar to the control L4440 fed cells (Fig.2.4A). SRs in these cells emanate from the basal body unit and extend from the posterior pole of the cell towards the anterior pole in a highly ordered pattern (Fig.2.15B, 2.15D). In contrast, ICL1e subfamily depleted cells show a very dramatic phenotype. In these cells, basal body rows misaligned on the surface of the cell are accompanied by the disordered organization of SRs (Fig.2.15E, 2.15F). Strikingly, in some areas of misaligned basal body rows, SRs point towards the posterior pole of the cell instead of the anterior pole of the cell (white arrows, Fig.2.15F). Misaligned basal body rows and disordered SRs organization are observed all over the cell surface except the oral groove area, suggesting that both the dorsal and ventral surfaces of the cells are similarly affected. All the data suggest that ICL1e subfamily protein may play an important role perhaps in the same pathways as SRs in maintaining the highly ordered organization of basal body units of the *Paramecium* cell.

Discussion

The motivation for this study was to understand how the SRs of *P. tetraurelia* help to maintain basal bodies in their rows that run from pole to pole of the cell. The SR emanates from 9 o'clock at the proximal end of the basal body and stretches upward toward the surface (Fig.2.2B, 2.2C). The SR then forms a straight line from the basal body toward the anterior of the cell past several more basal bodies. As shown in Allen (1971), two or three SRs can be found in parallel in a ridge of the cortical unit but untouching as they course toward the anterior. We identified a severe misalignment of these basal body rows associated with severe misalignment of SRs from their straight lines with the depletion of MKS3 ciliopathy gene (Picariello et al., 2014). The cell surface and its cortical units that harbor one or two basal bodies also were distorted from their expected regular hexagonal shape. By serendipity, the first *SR* gene that we targeted by RNAi was *SR7a*. We now know that this gene is identical to its partner (ohnolog *SR7b*) in its Paralog Group7 and together they form the Structural Group 2. The depletion of their mRNA led to consistent and dramatic basal body and SR row distortions. We subsequently identified the rest of the *SR* genes, which were not completely annotated, and found the relationships of these genes in Paralog and Structural Groups (more below). These findings provoked the question of how the phenotype of the depletions of Structural Groups developed. Do the SRs, which attach to the basal bodies, also attach to the structures under the cell body membrane like the ICL to maintain basal body rows? Are there reciprocal relationships, i.e. will disruption of the ICL cause changes in the organization of the SRs because attachment of the SRs to these cytoskeletal structures maintains the SRs organization along its long track toward the anterior? Do SR Structural

Group depletions affect the structure of the ICL because the SRs are an integral component in the pattern of the cortical cytoskeleton? We knew from a previous study using proteomics that isolation of the ICL included 25 of the 30 SR proteins that we identified (Gogendeau et al., 2008) although with different gene names because we did in this study a thorough annotation of the SR genes not previously known. However, there was no indication of how these large structures, SR and ICL, interacted. We show here that disruption of the SRs (Structural Group) leads to large distortions of the ICL units suggesting that the SR and ICL physically interact to anchor the basal body in its row and maintain the shape of the cortical unit. These large structures seem to be interacting with each other to make a straight SR row and highly repeating surface of the *Paramecium* from the basal bodies to the highly complex cortex (Fig.2.2).

In the course of these studies, we found a large number of *SR* genes (30) that appear to code for SR proteins, and some similar genes (*SRL*) that do not. Those that code for the SR proteins have in common the SF-assemblin sequences of the solitary *Chlamydomonas* SR protein. These SR genes can be sorted into 13 Paralog Groups that reflect the three whole genome duplications in *Paramecium.* More importantly, the Paralog Groups can be combined into what we call Structural Groups based upon their potential to have coiled coil structures in very particular positions in their proteins. There are five Structural Groups (Fig.2.1) whose members related by nucleic acid and amino acid sequences and the potential for coiled coil domains in the same part of their proteins. In addition, we consider a functional definition of the Structural Group: when all of the members of the Group are depleted in mRNA, the SR has a distorted and shorter shape, and apparently cannot interact with the ICL or other cortex components to maintain its

straight row and basal body rows. Depletion of its Paralog Groups does not have this effect. There is complete congruity among the phylogeny, primary amino acid sequences responsible for coiled coil domains, and the association with distorted surfaces (Table 2.5). Depletion of Paralog Groups (that alone do not form Structural Groups) by RNAi has no such effect, only depletion of entire Structural Groups.

The SR is composed of thirty proteins having the domains of the SF-assemblin protein:

Paramecium tetraurelia has a collection of large number of genes (~ 40,000) due to three successive whole genome duplication (Aury et al., 2006). The presence of the large number of *SR* genes and phylogeny suggests that these thirty genes are the result of these duplications. Our N-FLAG epitope tagging data for representatives of the thirteen SR Paralog Groups suggest that only the proteins having the characteristic domains of SF-assemblin protein localize in the SRs. SRs staining (using anti-SR antibody) in the wild-type cells and FLAG epitope staining in the FLAG-SR expressing cells demonstrate that FLAG-SR proteins are found throughout the SRs. Other similar proteins but which lack the domains of SF-assemblin has a different location than SRs. Collectively, our BLAST search using *Chlamydomonas* SF-assemblin in ParameciumDB, FLAG epitope tagging work of thirteen representative members from thirteen Paralog Groups, suggest that all thirty SR proteins, ranging 28-32 kDa molecular mass, are part of the SR structure. Moreover, identification of all thirty SR proteins by mass spectrometry after isolation of SRs also strengthens the idea that all the thirty proteins are the part of the SRs (Table 2.4).

Proteins from all five SR Structural Groups are required to maintain SR structure:

We assigned all thirty *SR* genes to five Structural Groups in part based on their predicted ability to form coiled coil domain. The coiled coil domain is a structural motif in proteins, thought to facilitate protein-protein interactions (Mason and Arndt, 2004; Truebestein and Leonard, 2016). These domains are thought to be involved in many important biological functions such as determining the architecture of the organelles (e.g., golgi, centrioles), acting as a molecular spacer to control the biological function of proteins, movement of the vesicles, molecular ruler of the enzymatic activity and many more (Mason and Arndt, 2004; Truebestein and Leonard, 2016). The length and position of the coiled coil domains in the SR protein sequences are similar within a Structural Group but vary between Structural Groups, suggesting that coiled coil domain of proteins within a Structural Group may have the same function of protein-protein interaction.

We used RNAi feeding to explore the function the SR proteins in maintaining the normal appearance of the SRs. With the depletion of a single Structural Group, the SRs become short and the striation pattern changes dramatically. The RNAi data suggest that all five SR Structural Groups contribute to the overall SR appearance. Interestingly, depletion of one gene for an SR protein or the members of one Paralog Group within a Structural Group does not cause the change in length and abnormal appearance except for Paralog Groups 7 which is synonymous with the Structural Groups 2 (Fig.2.1). The Structural Group 2 consists of two ohnolog genes (*SR7a*and *SR7b*). The nucleotide sequences of two genes are very similar and their amino acid sequences are identical. RNAi off-target analysis showed RNAi construct for *SR7a* affects both *SR7a* and *SR7b*.

Therefore, depletion of *SR7a* also leads to the depletion of *SR7b*, Paralog Group 7 and Structural Group 2.

Ultra structure studies of isolated rootlets from control cells, SR Paralog Group depleted cells or SR Structural Group depleted cells clearly suggest that all five Structural Groups of SR proteins are important for the formation of SRs with normal length and striations. In the study (Lechtreck, 1998) of SF-assemblin protein, the sole constituent of *Chlamydomonas* SR describes various mutations of the SF-assemblin changes the solubility, striation pattern and overall appearance of the SRs. We also observe changes in the major striation periodicity or the absence of striation pattern in response to the depletion of any of the five SR Structural Groups. The proteins from the *Chlaymdomonas'* single *SR* gene can be predicted to overlap and interact through their coiled coil domains and overlap with the N-terminal and C-terminal portions of the protein to form the fiber and the striations. Our data (immunofluorescence and transmission electron micrograph of isolated SRs) clearly demonstrate all five Structural Groups are required to interact with each other to form an SR with normal length and striation. How one Structural Group interacts with other Structural Group remains to be answered. The resolution (~200nm) of immunofluorescence data of N-FLAG tagged SR proteins (representative members of 13 Paralog Groups which further categorize into 5 Structural Group) is not enough to provide the clue how the members of 5 Structural Groups interact with each other. We used the data obtained from the ultra structure studies of isolated SRs (striation pattern) to create a model describing how possibly SR protein members of 5 different Structural Groups interact with each other. The model is described in Chapter 4 in details.

Collectively, these data uncover some novel findings for the proteins of the SRs. Firstly, SR protein members within a Structural Group seem to have the ability to replace each other, meaning depletion of one member in a Structural Group can be adjusted by other members of the Structural Group, which have the same coiled coil structure. Secondly, members of all five Structural Groups contribute to the characteristic striation pattern in the complex striated rootlet structure.

SR depletion does not appear to affect the positioning of the three rootlets on the basal body:

In ciliates, the core structure of the basal body consists of nine sets of microtubule triplets that align in a cylindrical shape (Frankel 2000). In *Paramecium*, the basal body structure has the same microtubule triplet core structure and attaches to a set of three rootlets. One of the rootlets is the SR. The other two rootlets are the microtubule ribbons called the PCR and the TR (Tassin et al., 2016). The basal bodies along with their three asymmetrical rootlets align in a highly organized straight line from the posterior to the anterior pole of the cell on both dorsal and ventral surfaces (Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003). Rootlets attach to the basal bodies at precise locations in regard to the microtubule triplets. The SR, PCR and TR rootlets attach to the microtubule triple 5, 9 and 4 of the basal body respectively in *Paramecium*, forming reproducible angles between the rootlets (Jerka-Dziadosz et al., 2013). To examine whether the SR, PCR and TR attach to the basal bodies at the normal sites in SR Paralog Group or Structural Group depleted cells, we examined the angles between these rootlets (the angles between PCR and TR; the angles between PCR and SR). We chose to examine control, SR Paralog Group 2, SR Structural Group1 (Paralog Group: SR 1) and SR Structural Group 2

(Paralog Group: SR 7) depleted cells. The angles between the PCR and TR remain constant (~109°) in control, SR Paralog Group 2, SR Structural Group 1 (Paralog Group: SR 1) or SR Structural Group 2 (Paralog Group: SR 7) depleted cells. The angles between SR and PCR also remain constant (~136°) in all the cells. The constant angles between rootlets in different cells may be interpreted as follows: Depletion of SR Structural Groups cause basal body row misalignment but does not seem to affect the integrity of the basal body structure.

Full length of SRs is required to interact with other ICL centrins to maintain the straight basal body rows:

In green flagellates, striated microtubule associated fibers have assumed to play role in anchoring flagella (Hyams and Borisy, 1975; Sleigh and Silvester, 1983). A study in ciliate organism *Tetrahymena* describes a regulator protein DisA, is required to maintain length of SRs which eventually plays a role in basal body orientation in multi-ciliary rows (Galati et al., 2014). Without DisA, the SR shortens and the physical stress of ciliary beating throws the basal bodies out of alignment. In mammalian system, the SR is a one of the largest cytoskeleton structures, originating from the basal body or the base of the cilium and extending proximally toward the cell nucleus (Yang et al., 2002; Yang et al., 2005). Like Chlamydomonas SF-assemblin protein, the SR in mammalian system is made of one single protein designated as rootletin. Rootletin is a 220kDa protein, has a N-terminal globular head domain and C-terminal coiled coil domain. The main function of the SR in the mammalian system is to provide the structural stability of the cilium (Yang et al., 2002; Yang et al., 2005).

In *Paramecium*, the SR is the most prominent structure among the three rootlets that are associated with the basal body unit (Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003). In a dividing cell, SRs show remarkable regression and re-growth which couple with the duplication of the basal body. Unlike basal bodies in mammalian cells, basal body duplication in *Paramecium* occurs near the mother basal body, which is anchored to the cell surface. Rootlets have been assumed to act as a scaffold to define the position of the new daughter basal body. Ultra structure studies of the basal body duplication, regression and re-growth of SRs suggest a role of SRs in guiding the daughter basal body in growing and dividing cells (Iftode et al., 1996; Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003; Tassin et al., 2016).

Our work shows that there is a role for SRs in interphase cells in maintaining straight basal body rows. Depletion of any of the Structural Group results in misalignment of the basal body rows. Misalignment of basal body rows can occur anywhere over the cell surface (dorsal surface, mid-dorsal region, anterior or posterior pole region and ventral surface) except the oral groove area. However, our immunofluorescence data demonstrates that basal body row misalignment shows spatial heterogeneity in a single cell. Moreover, the area of misalignment of the basal body rows varies from cell to cell depleted of the same Structural Group (e.g., Structural Group 2). The ultra structure study (Hufnagel, 1969) by Hufnagel of the *Paramecium* cortex demonstrates cortical units in a single cell can have variable characteristics. There is a similarity between the structural organization (basal body unit at the center of the cortical territory, basal body associated microtubule rootlets, SR, epiplasm, alveolar membrane and parasomal sac) of the different cortical territories in a single cell, but each cortical

territory may show variable characteristics that include the number of basal body units, length of SR, number of microtubule rootlets and number of parasomal sacs. The heterogeneous nature of the cortical unit that harbors the basal body unit depends on the surrounding environment of the cortical territory. All these characteristics of the cortical unit territories suggest that RNAi of SR structural Group may have a variable effect on the different surface areas of the *Paramecium* cortex.

The infraciliary lattice (ICL) is the inner most cytoskeleton structure in cortex of *Paramecium*. The ICL contractile network of irregular shaped polygons forms a mesh that is the innermost layer of each cortical unit. It sits beneath the proximal end of the basal bodies (Allen, 1971; Garreau de Loubresse et al., 1988; Garreau de Loubresse et al., 1991; Gogendeau et al., 2008). The ICL is composed of 35 centrin proteins and 3 large Sfi1P like centrin binding protein. Among the three centrin binding proteins, the protein encoded by *PtCenBP1* plays an important role in the formation of the backbone of the ICL network, whereas centrin proteins encoded by *PtCenBP2/PtCenBP2* may function in attaching the ICL network to the cell cortex. Silencing of the centrin binding proteins suggests none of these proteins have a function in maintaining the ciliary basal body rows or locking the mature basal body in the cell cortex (Gogendeau et al., 2007; Gogendeau et al., 2008). The centrin proteins that are the constituents of the ICL network can be divided into 10 subfamilies e.g., ICL1a, ICL1e, ICL3a, ICL3b, ICL5, ICL7, ICL8, ICL9, ICL10 and ICL11 (Gogendeau et al., 2008; Aubusson-Fleury et al. 2017). All of these are found in the ICL network and have specific functions related to the network, with the exception of subfamily ICL1e.The ICL1e subfamily consists of seven ICL-centrin proteins(Ptcen_icl1e, Ptcen_icl1g, Ptcen8, Ptcen10, Ptcen12, Ptcen15 and Ptcen18); all

the seven genes code for proteins with highly similar sequences (the number of amino acid ranges from 174-178).The ICL 1e subfamily proteins appear to have multiple locations; PtCenICl1e is found at the ICL network, contractile vacuole, and basal body (Gogendeau et al., 2008). Silencing of all of the ICL1e subfamily genes by using an RNAi construct containing the entire sequence of *Ptcen_icl1e* (Gogendeau et al., 2008; Aubusson-Fleuryet al. 2017) results in the disruption of the ICL network, misalignment of basal body rows and disorganized oral apparatus. (The SR organization was not examined in these previous studies)These data suggest that the ICL1e subfamily may function to maintain physical link between the ICL network and other cortical organelles (Gogendeau et al., 2008).

A clue about the reciprocal interaction between the ICL networks and SRs comes from the analysis of our LC-MS/MS data of isolated SRs and LC-MS/MS data of isolated ICL (Gogendeau et al., 2008). In the previous study (Gogendeau et al., 2008) ICL was isolated from cell cortex by using lysis buffer, followed by using buffer containing urea and pelleted by centrifugation. We isolated SRs from the cell cortex by using a combination of buffers, mechanical forces (homogenization of the cells) and density gradient ultracentrifugation. Although the isolation procedures diff, they both likely isolate large cytoskeletal structures that overlap in their components. The proteomic analysis of the ICL identified the protein members of SRs (25 SRs among 30 SRs) (Gogendeau et al., 2008) and our proteomic analysis of SRs identified the protein members of ICL network (ICL1e subfamily proteins and centrin binding proteins). Interestingly, proteomic analysis of the isolated SRs from SR Structural Group depleted cells (SR Structural Group 1 or SR Structural Group 2) failed to identify any proteins

from ICL network. The presence of other proteins (α-tubulin, β-tubulin, heat shock proteins and ribosomal proteins) might indicate the contamination of the SRs isolation procedure because these proteins are also identified in proteomic analysis of the isolated SRs from SR Structural Group depleted cells. All these data strongly suggest there might be a reciprocal relationship exists between SRs and ICL mesh network.

In this study, to tease apart the possible reciprocal interactions between ICL network and SRs, we silenced separately SR Structural Groups 1 and 2 and examined the ICL network using an expressed protein FLAG-Ptcen15 to visualize the ICL network and the impact of depletion of SR Structural Groups. Subsequently in non-expressing cells, we silenced the ICL1e subfamily and observed the SR organization. We also silenced two additional ICL centrin genes as a control since that were not implicated in maintaining physical link between the ICL network and other cortical organelles reported by Gogendeau et al., 2008. (ICL9 and ICL10 subfamilies have minor effects on the ICL network but do not have any effect on basal body row organization upon silencing (Gogendeau et al., 2008). We speculated depletion of ICL1e subfamily should affect the ordered organization of the SRs and misalign basal body rows if a relationship of reciprocal interactions exists between ICL network and SRs.

To test our idea of reciprocal interactions, we silenced the ICL1e subfamily by RNAi using the entire gene sequence of *Ptcen_icl1e* to co-silence all seven genes of the ICL1e subfamily and examined the effects on the SRs and basal body rows. Our RT-PCR analysis of ICL1e subfamily RNAi data using *Ptcen12* or *Ptcen15* specific primers clearly demonstrates *Ptcen_icl1e* RNAi construct reduces the ICL1e subfamily transcripts. Our immunofluorescence data clearly demonstrate that the depletion of the

ICL1e subfamily results in misaligned basal body rows along with highly disordered SRs all over the cell surface. Strikingly, in some cases ICL1e subfamily depletion severely changes the direction of SRs, meaning SRs point toward the posterior pole of the cell instead of anterior pole. Moreover, depletion of other ICL subfamily members ((ICL 9 subfamily (*Ptcen_icl9a*geneRNAi construct) or ICL10 subfamily (*Ptcen_icl10a*gene RNAi construct)) appears to have no effect on the alignment of basal body rows or the highly ordered SRs.

All of the test results for reciprocal interactions are consistent with the hypothesis that the ICL1e subfamily proteins are the sites of interaction with SRs to maintain straight basal body rows and the highly ordered organization of SRs from the posterior pole towards the anterior pole of the cell cortex.

Silencing of SR Structural Groups profoundly disrupted the ICL organization as well as SR organization and structure; conversely, depletion of ICL1e dramatically affects both the ICL and the highly ordered SRs, suggesting that there is a reciprocal relationship between SRs and ICL network, possibly through physical protein-protein interactions. The location of the ICL1e-SR interaction could be in the ICL lattice or at the basal body itself as both the SR and ICL1e members can be found in both locations. The SR and ICL1e proteins can be found together in the SR isolation density fraction, but many other proteins are similarly found there. Therefore, we have only co-isolation and correlative RNAi data of interactions.

How the SR depletion affects the shape of the cortical unit:

Just beneath the plasma membrane that covers the outer surface of the cortical unit is the epiplasm layer. This layer seems to line the cortical unit, and thus appears to be in territories that coincide with each cortical. In the center of the territory is a hole through which the cilium emerges (Nahonet al., 1993; Aubusson-Fleury et al., 2013). In *Paramecium*, epiplasm consists of large family of 51 proteins, which are categorized as rim epiplasmin, core epiplasmin, or ring and basal body epiplasmin depending on their location in the interphase cell. The core epiplasmin appears to play role in building the new epiplasmin unit during cell division (Damaj et al., 2009; Aubusson-Fleury et al., 2013). Silencing of the core epiplasmin transcripts impairs with cell division process of the cell and eventually affects the cell shape and morphology due to impaired cell division process. The rim epiplasmin appears to play a role in the segregation of the new epiplasmin unit and silencing of this group results in partial aggregation of epiplasm units during cell division. The ring and basal body epiplasm does not appear to play important role in basal body anchoring and silencing of this group of epiplasmin has almost no effect on the cell shape and growth (Damaj et al., 2009; Aubusson-Fleury et al., 2013). A new finding from our study is that there appears to be a role for SRs in shaping the morphology of the epiplasm territory.

We believe that the effect of SR depletion on the epiplasm territories is not an indication of an SR-epiplasmin interaction for several reasons. Because SR depletion severely disrupts the ICL and cortical unit shape as shown by cortical ridge antibodies, we expected the epiplasm territories to likewise be disrupted because of the distortion of the ICL. In previous studies, RNAi for each of the epiplasmin groups does not appear to have much effect on the straight basal body rows from the anterior pole towards the

posterior pole of the cell (Damaj et al., 2009; Aubusson-Fleury et al., 2013), compared to our RNAi results in which the misalignment of basal body rows is the signature phenotype of SR Structural Group depletion. Last, our LC-MS/MS did not identify any epiplasmin proteins that co-isolated in the SRs preparation, suggesting that there is not the same interaction of SR proteins and epiplasmins as there is with ICL1e proteins.

Fig.2.2B demonstrates the arrangement of cortical unit to show the ridge with 3 SRs coursing through it, the basal body resting on the ICL in the innermost layer of the cortical unit. Fig.2.2A shows the level of interaction between ICL and the SR that we propose is necessary to maintain the ridge and epiplasm territory and the direction of basal bodies and SRs. We postulate that the ICL and SR interact near to where the SR attaches to the basal body. From there the SR can be followed upward to where it enters the ridge. There are runs in parallel to other SRs from other cortical units, but these SRs do not seem to touch. The failure of the SR-ICL interaction allows the SR to lose its way, and not keep the straight path toward the anterior. The rotation of basal bodies out of line is a consequence of these changes. A model describing the possible interaction of SR and ICL network outlines in Chapter 4.

Swimming behavior changes in SR Structural Group depleted cells:

Immunostaining of the cells with anti-α tubulin demonstrates that all the cells (control or SR Paralog Group depleted cells or SR Structural Group depleted cells) have cilia despite having misaligned basal body rows in the SR Structural Group depleted cells. This result suggests depletion of SR Structural Groups does not affect the biogenesis of cilia, and, therefore, the cells should move their cilia and swim. The *Paramecium* cell swims by movement of their thousands of cilia. Each cilium goes

through a beat cycle consisting of a power (or effective) stroke and a recovery (or return) stroke (Kung et al., 1975; Satir et al., 2014). In a power stroke, which allows the cell to move forward, the cilium makes obliquely backward movement to the longitudinal axis of the cell body with force against the surrounding medium. The power stroke is followed by a return stroke in which the cilium makes a counter-clockwise rotation parallel to the cell surface for returning to its regional curved position (Kung et al., 1975; Satir et al., 2014).

Given that depletion of Structural Group SRs leads to misaligned basal body and SR rows, we predicted that the swimming pattern of these cells would deviate from the normal straight swimming of control cells. Indeed, the RNAi treated cells for Structural Groups but not for individual *SR* genes or Paralog Groups showed swirly and inefficient swimming. Analyzing the tracks of these cells further would not shed any more light on our goal of understanding how SRs and the cortex interact, but the abnormal tracks nicely show that there is a behavioral consequence of having basal body and SR rows out of alignment. These tracks also demonstrate that the cells' cilia beat allowing them to swim and perhaps throw the basal body rows out of alignment.

Concluding Remarks:

In this study, we identified in total thirty *SR* genes for proteins that belong to five Structural Groups based on their primary sequence and expected ability to form coiled coil domains. From examination of representatives of each Structural Group, it appears that all of these thirty proteins are throughout the length of SR without showing any domain specific localization for a particular SR Structural Group. All five Structural Groups are required in order to assemble normal SRs in *Paramecium tetraurelia*. In the

absence of one Structural Group, the length of the SR shortens and the striation pattern changes dramatically. As a result, abnormal SRs appear to lose the interaction with other cytoskeleton structures such as ICL network complex, which eventually results in misaligned basal body rows and altered swimming behavior. It is reasonable to postulate that protein members of the ICL1e subfamily and SRs are in a reciprocal relationship to maintain the straight basal body row and the highly ordered organization of the SRs all over the cell surface. Possibly the ICL1e subfamily proteins and SRs work together to maintain the straight alignment of basal body rows by locking the basal body units at cell cortex. It is also possible that other cortical cytoskeleton structures such as epiplasm that respond to the distortion of the ICL show the effects of SRs depletion but are not the cause of the basal body row distortions. Perhaps it is the ciliary beating that causes the basal bodies with their SRs to move out of alignment when the interactions among the SRs, ICL and epiplasm are not in place to resist this movement.

Our model for the interaction of the SRs and ICL to maintain basal body rows is shown in detail in Chapter 4.

Figures and tables

Fig.2.1: The phylogenetic relationships among the *SR* genes in *Paramecium*.

Red circles in panel A show the five structural groups based on their ability to form coiled coil domain. Panel B shows the position of the coiled coil domains in the five Structural Groups. The alignment of the coiled coil regions at the amino acid level is shown in supplement Fig.2.S1.

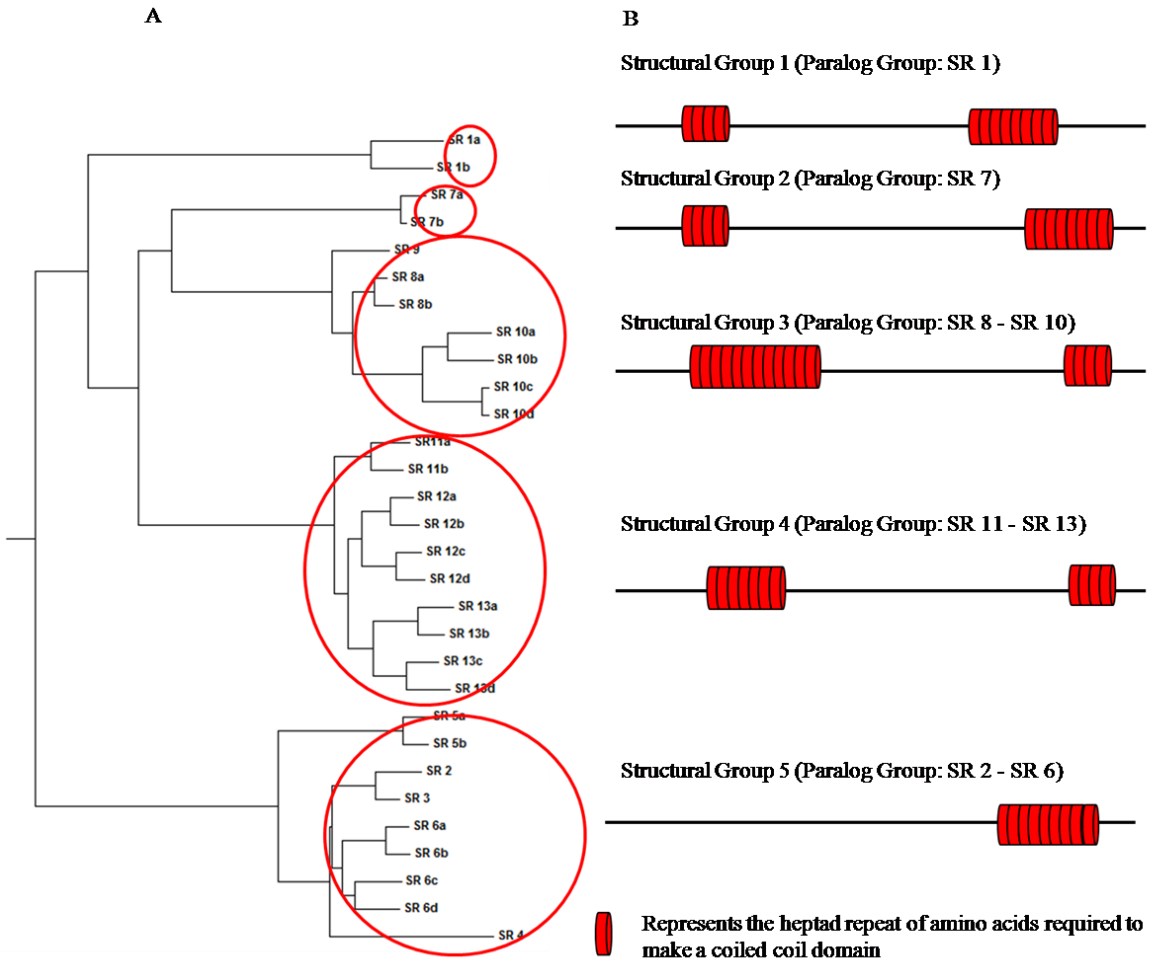
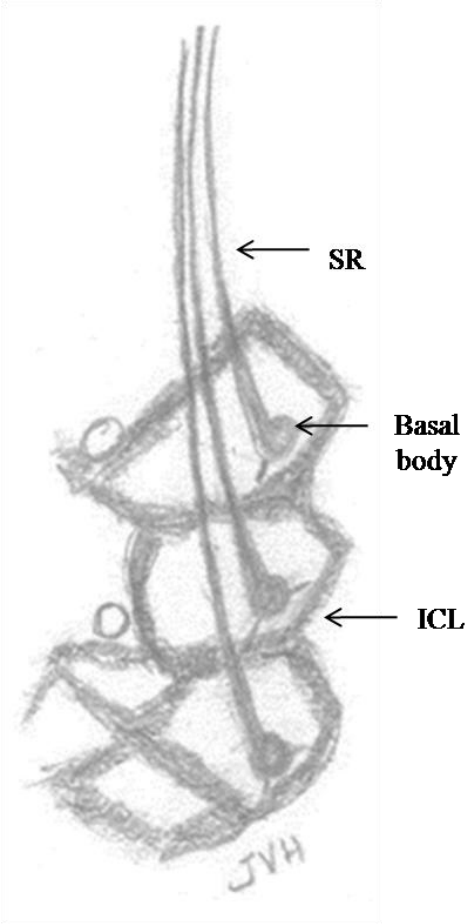
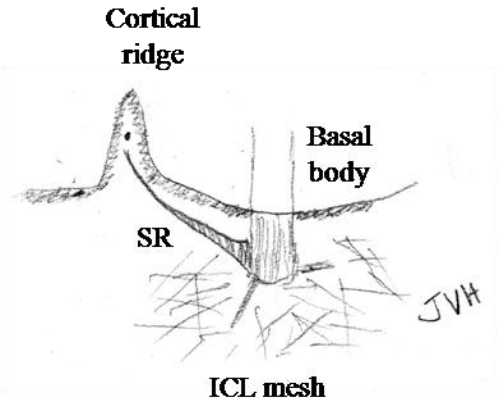

A
SR 1a
SR 1b
SR 7a
SR 7b
SR 9
SR 8a
SR 8b
SR 10a
SR 10b
SR 10c
SR 10d
SR11a
SR 11b
SR 12a
SR 12b
SR 12c
SR 12d
SR 13a
SR 13b
SR 13c
SR 13d
SR 5a
SR 5b
SR 2
SR 3
SR 6a
SR 6b
SR 6c
SR 6d
SR 4
B
Structural Group 1 (Paralog Group: SR 1)
Structural Group 2 (Paralog Group: SR 7)
Structural Group 3 (Paralog Group: SR 8 - SR 10)
Structural Group 4 (Paralog Group: SR 11 - SR 13)
Structural Group 5 (Paralog Group: SR 2 - SR 6)
Represents the heptad repeat of amino acids required to make a coiled coil domain

Fig. 2.2: A highly organized *Paramecium tetraurelia* cell cortex.

Panel 2A shows the organization of cytoskeleton structures (SRs, basal body units and ICL) in *P. tetraurelia*. The basal body structure is associated with a set of three rootlets. Two of rootlets are microtubule based rootlets: PCR project toward the posterior pole of the cell and TR projects laterally towards the adjacent rows of basal bodies. SR is the largest rootlet that originates from the basal body and extends several anterior basal body units. ICL is the innermost cytoskeleton contractile network. All these cytoskeleton structures along with basal body unit and its associated rootlets form the highly organized cortex of the *Paramecium* which is characterized by straight rows of basal body units from the posterior pole towards the anterior pole of the cell. Panel B shows how SR runs through the ridge of the cortical unit to maintain the straight basal body rows and the highly organized cell surface of the *Paramecium*.


SR
Basal body
ICL
JVH

A


Cortical ridge
Basal body
SR
JVH
ICL mesh

B

Fig.2.3: SR proteins that have the conserved domains of the *Chlamydomonas* SF-assemblin protein are confirmed to be in the SR.

Cells expressing only the FLAG epitope (control) or a representative FLAG-SR were immunostained with anti-centrin (red) and anti-FLAG (green). Panel shows the side view images of control, SR1a, SR2, SR3, SR4, SR5a, SR6b, SR7a, SR8a, SR9, SR10c, SR11a, SR12b, SR13c or wild type cells (no exogenous expression). The basal bodies are red and the SRs are green. In the wild type image, basal bodies are green (anti-1D5) and SRs are red (anti-SR). All the representative members of the SR proteins (green) are located throughout the SRs. Scale bars are 3μm.

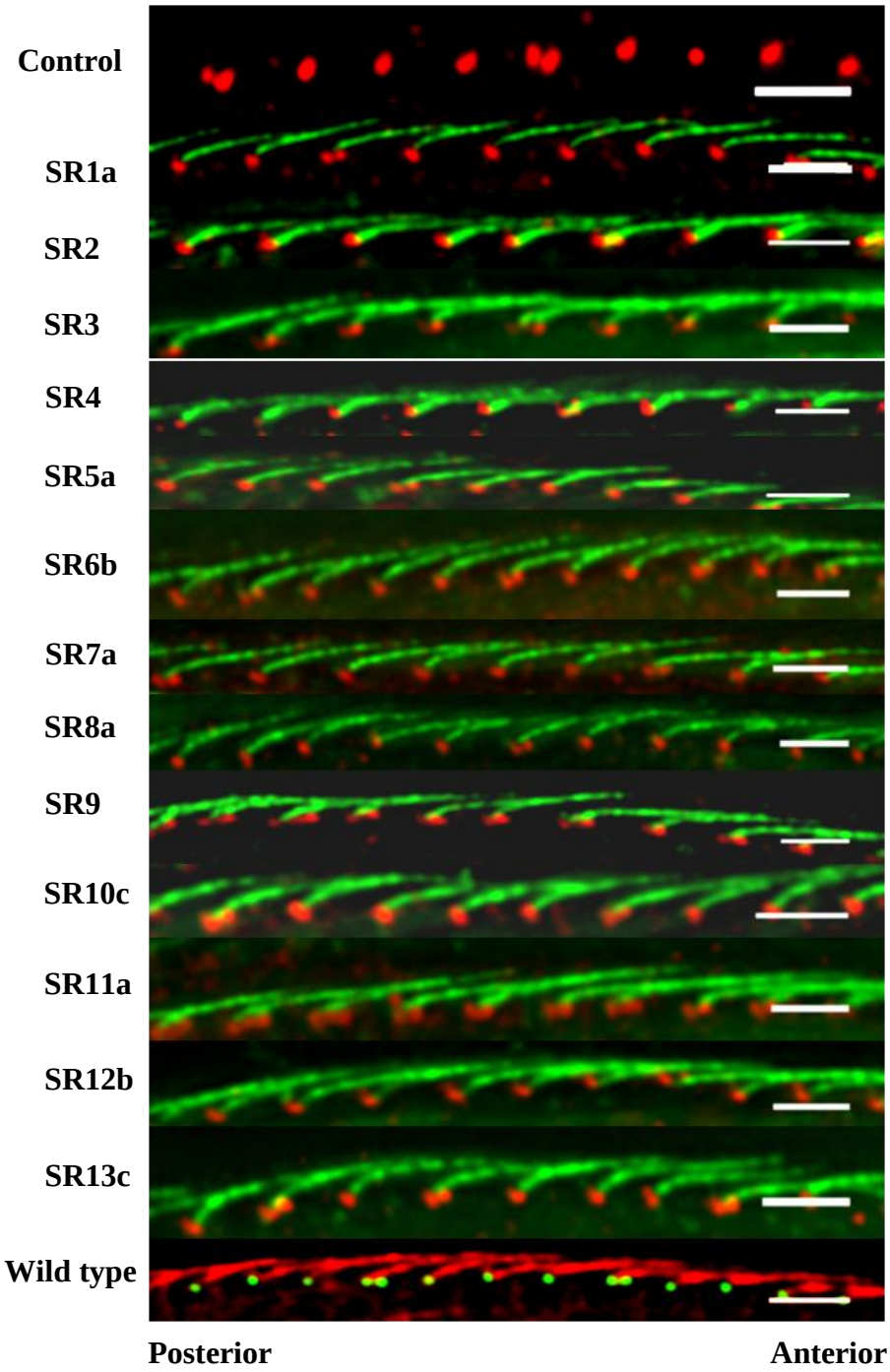
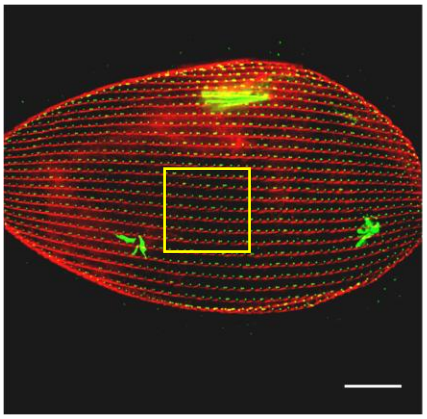

Control
SR1a
SR2
SR3
SR4
SR5a
SR6b
SR7a
SR8a
SR9
SR10c
SR11a
SR12b
SR13c
Wild type
Posterior
Anterior

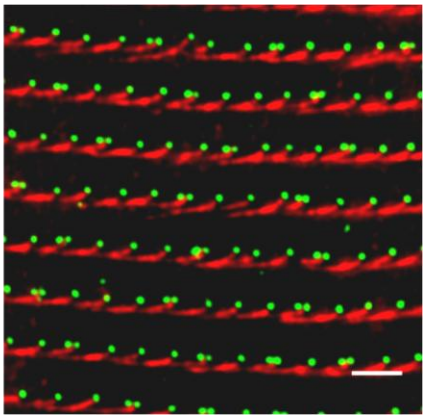
Fig.2.4: SR Structural Group depletion can lead to basal body row misalignment and SR abnormal appearance.

Panels A and C show control and Structural Group 2 (Paralog Group: SR 7) depleted cells respectively. The yellow box in each image is enlarged (B and D) to highlight basal body rows. The basal bodies are green (anti-ID5) and SRs are red (anti-SR). All images are of the dorsal surface. Panels A and B show straight rows of basal bodies as well as SRs that extend from the posterior pole towards the anterior pole. Panels C and D show the severely misaligned basal body rows as well as SRs with abnormal appearance (shorter and not pointed towards the anterior pole). Misalignment occurs over the entire cell surface. Scale bars are 15μm (A and C) and 3μm (B and D).

Control cell

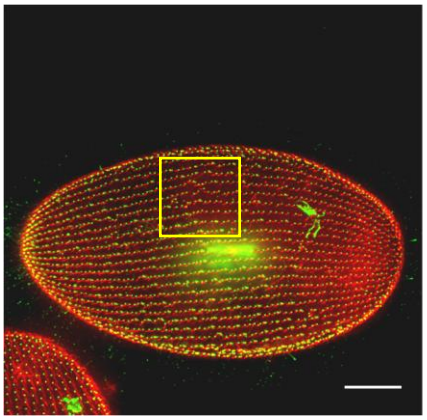


A

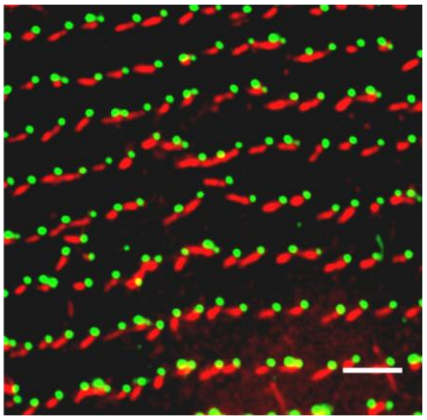


B

SR Structural Group 2 (SR 7) depleted cell



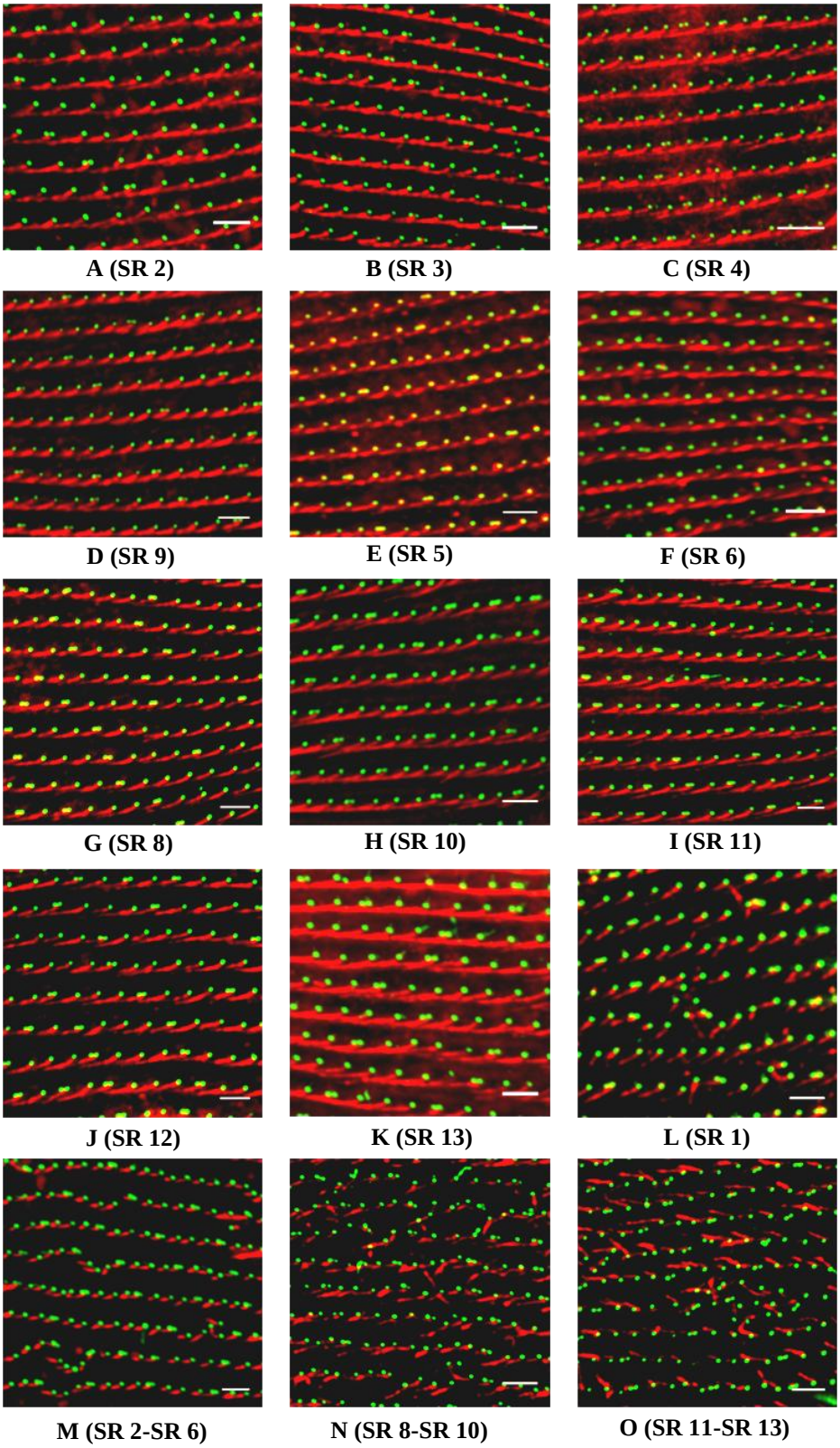
C



D

Fig.2.5: Basal body row alignment and SRs appearance in cells with different SR Paralog Group or Structural Group depleted cells.

Panels A-K show the representative images of the basal body rows (green) and SRs (red) in the SR Paralog Group SR 2, SR 3, SR 4, SR 9, SR 5, SR 8, SR 10, SR 11, SR 12 or SR 13 depleted cells. Panels L-O show the representative images of the basal body rows (green) and SRs (red) in the SR Structural Group 1 (Paralog Group: SR 1), Structural Group 5 (Paralog Group: SR 2-SR 6), Structural Group 4 (Paralog Group: SR 8-SR 10) or Structural Group 3 (Paralog Group: SR 11-SR 13) depleted cells. Depletion of Paralog Group (Panels A-K) does not affect basal body row alignment and the highly ordered organization of SRs appearance. Depletion of Structural Groups (Panels L-O) results in misaligned basal body rows with abnormal SRs appearance. Scale bars are 3μm. Images of the whole cell are available in the supplement (Fig.S4). Panels G, H and I have been rotated for easier comparison.



A (SR 2) **B (SR 3)** **C (SR 4)**

D (SR 9) **E (SR 5)** **F (SR 6)**

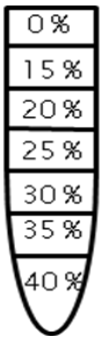
G (SR 8) **H (SR 10)** **I (SR 11)**

J (SR 12) **K (SR 13)** **L (SR 1)**

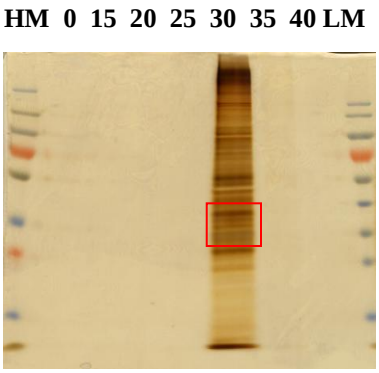
M (SR 2-SR 6) **N (SR 8-SR 10)** **O (SR 11-SR 13)**

Fig.2.6: Isolation of SRs from FLAG-SR8a expressing cells using Optiprep density gradients.

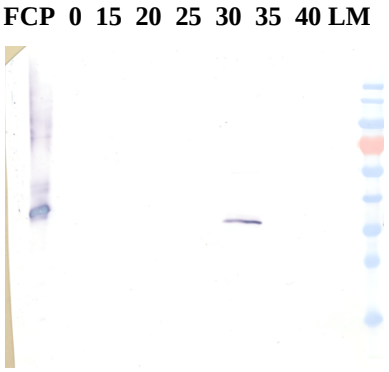
Panel A shows the different percentages of Optiprep in the step density gradient. Panel B shows all visible proteins concentrate in the 30% Optiprep fraction in a silver stained gel. Panel C shows expressed FLAG-SR8a proteins present in the 30% Optiprep fraction at the expected area of molecular mass (~33 kDa). The blot was treated with anti-FLAG antibody as a primary antibody. Panel D shows the immunofluorescence image of a SR from the 30% Optiprep fraction embedded in agarose. Scale bar is 2μm.


0%
15%
20%
25%
30%
35%
40%

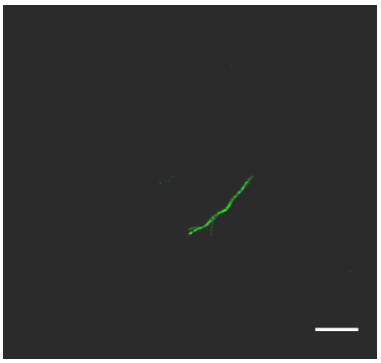
A


HM 0 15 20 25 30 35 40 LM

B


FCP 0 15 20 25 30 35 40 LM
Anti-Flag

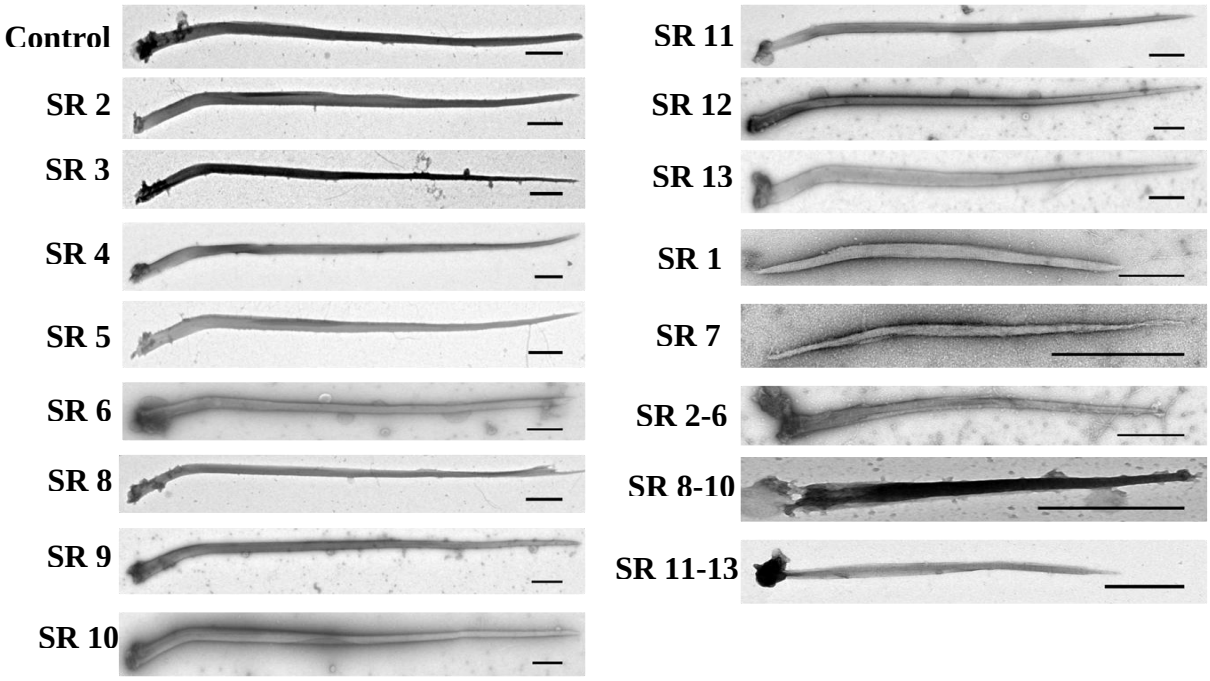
C



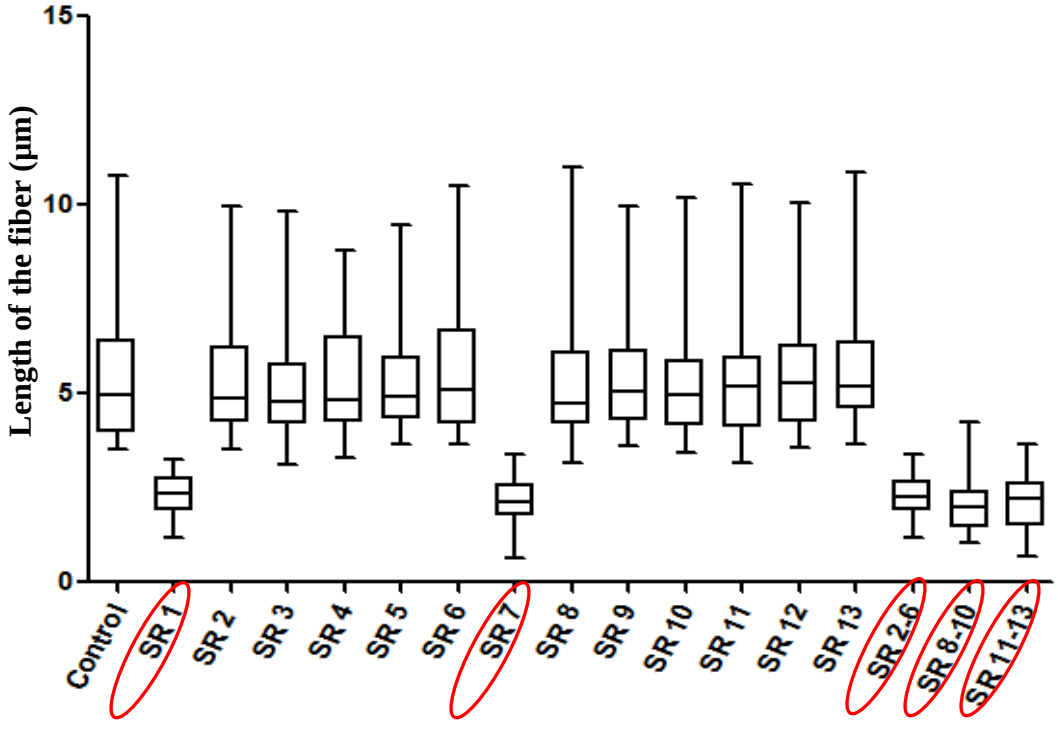
D

Fig.2.7: Depletion of SR Structural Group causes changes in the length of the SRs.

Panel A shows the electron micrograph of the SRs isolated from control, Paralog Group (SR 2, SR 3, SR 4, SR 5, SR 6, SR 8, SR 9, SR 10, SR 11, SR 12 and SR 13) or Structural Group ((Structural Group 1: SR 1), (Structural Group 2: SR 7), (Structural Group 3: SR 8-SR 10), (Structural Group 4: SR 11-SR 13) and (Structural Group 5: SR 2-SR 6)) depleted cells. Scale bars are 0.5μm. Panel B shows the length of the isolated SRs from Structural Group (SR 1), Structural Group 2 (SR 7), Structural Group 3 (SR 8-SR 10), Structural Group 4 (SR 11-SR 13) or Structural Group 5 (SR 2-SR 6) are significantly shorter than the control or all other SR Paralog Group depleted cells.



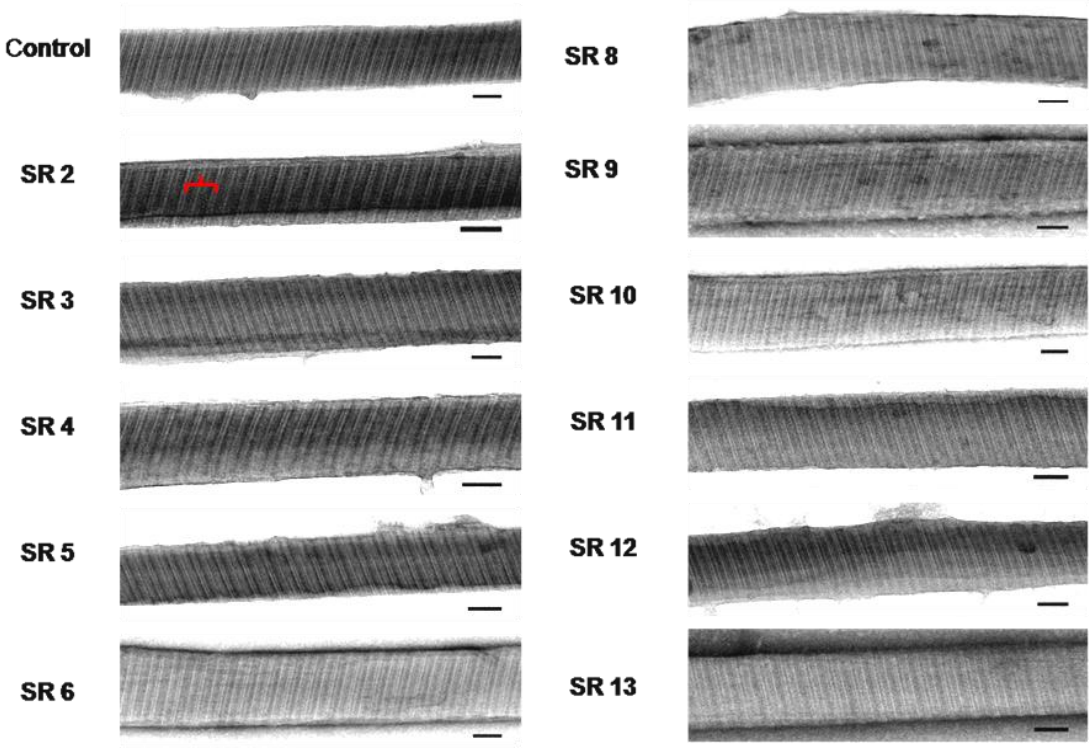
A (SRs isolated from different cells)



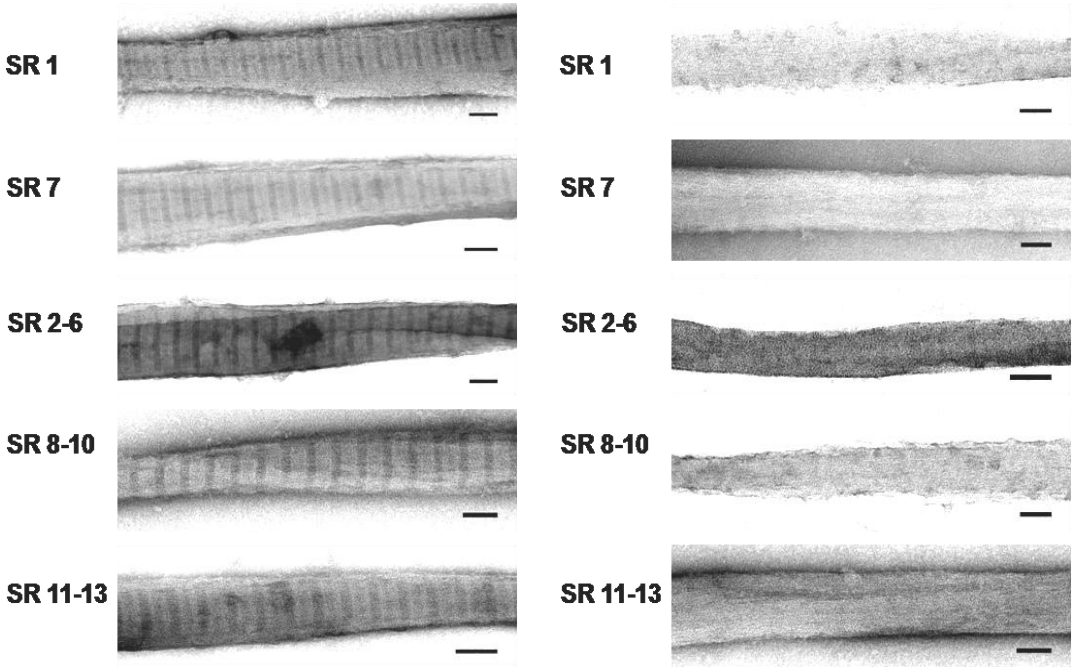
B (Length comparison of SRs isolated from different cells)

Fig.2.8: Depletion of SR Structural Group changes the striation pattern of the isolated SRs.

SRs in panel A show the striation pattern in the control and different Paralog Group (SR 2, SR 3, SR 4, SR 5, SR 6, SR 8, SR 9, SR 10, SR 11, SR 12 and SR 13) depleted cells (where SR appearance does not change). In these cells, major striations occur at 24-34nm intervals. SRs in panel B shows the striation pattern is missing in the SR Structural Group (Structural Group 1: SR 1), (Structural Group 2: SR 7), (Structural Group 3: SR 8-SR 10), (Structural Group 4: SR 11-SR 13) and (Structural Group 5: SR 2-SR 6) depleted cells. Scale bars are 50nm.


Control
SR 2
SR 3
SR 4
SR 5
SR 6
SR 8
SR 9
SR 10
SR 11
SR 12
SR 13

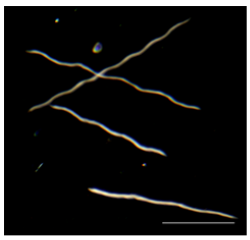
A


SR 1
SR 7
SR 2-6
SR 8-10
SR 11-13
SR 1
SR 7
SR 2-6
SR 8-10
SR 11-13

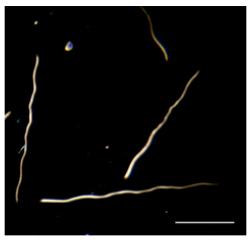
B

Fig.2.9: Changes in swimming behavior happens with the depletion of SR Structural Group.

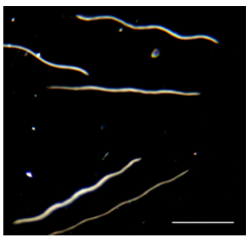
Panels show the swimming pattern in different cells taken by dark field microscopy. For all images, camera exposure time was 2 sec. Depletion of Paralog Group (SR 2, SR 3, SR 4, SR 5, SR 6, SR 8, SR 9, SR 10, SR 11, SR 12 or SR 13) do not affect swimming behavior of the cells as compared to control cells. Depletion of SR Structural Group ((Structural Group 1: SR 1), (Structural Group 2: SR 7), (Structural Group 3: SR 8-SR 10), (Structural Group 4: SR 11-SR 13) or (Structural Group 5: SR 2-SR 6)) result in very different swimming behavior of the cells as compared to control cells. Scale bars are 1mm.



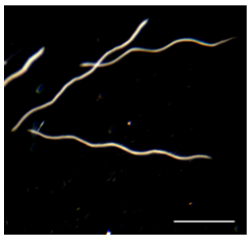
A (Control)



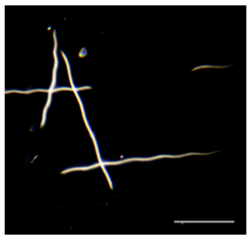
B (SR 2)



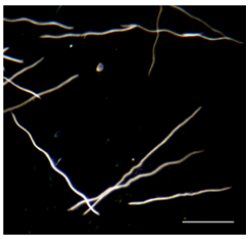
C (SR 3)



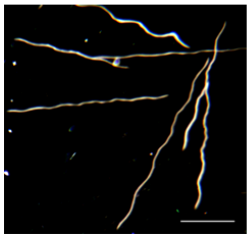
D (SR 4)



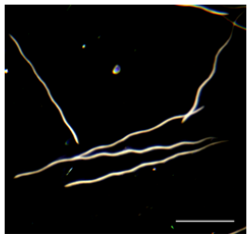
E (SR 5)



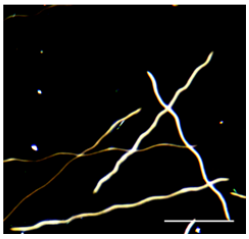
F (SR 6)



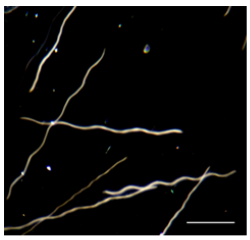
G (SR 8)



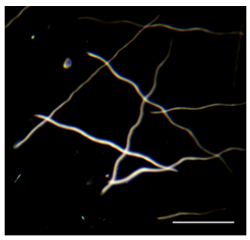
H (SR 9)



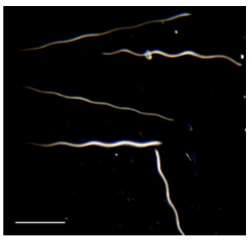
I (SR 10)



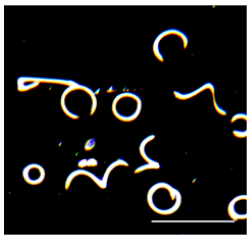
J (SR 11)



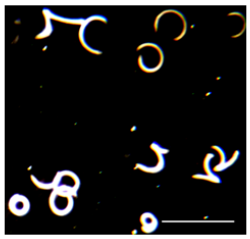
K (SR 12)



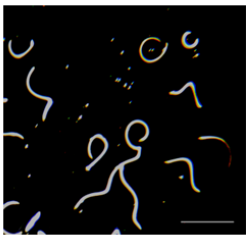
L (SR 13)



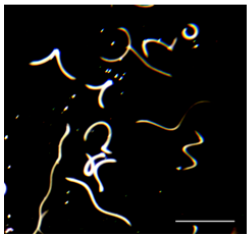
M (SR 1)



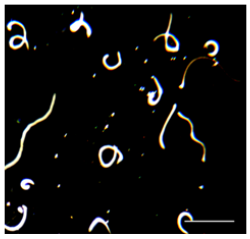
N (SR 7)



O (SR 8 - SR 10)



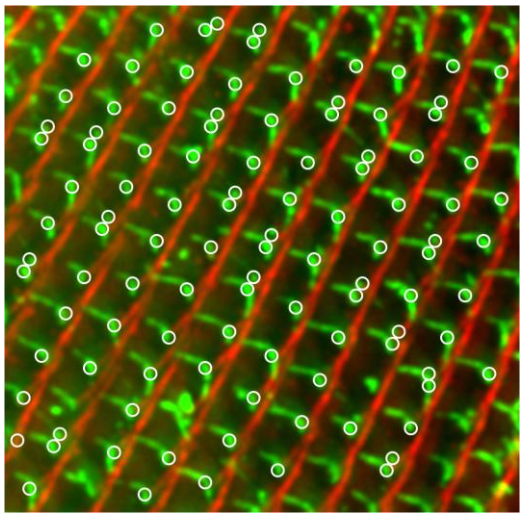
P (SR 11 - SR 13)



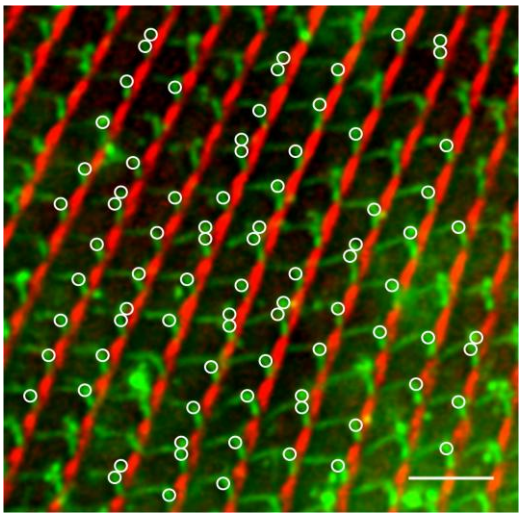
Q (SR 2 - SR 6)

Fig.2.10: Depletion of SR Paralog Group or SR Structural Group does not affect the angles between rootlets.

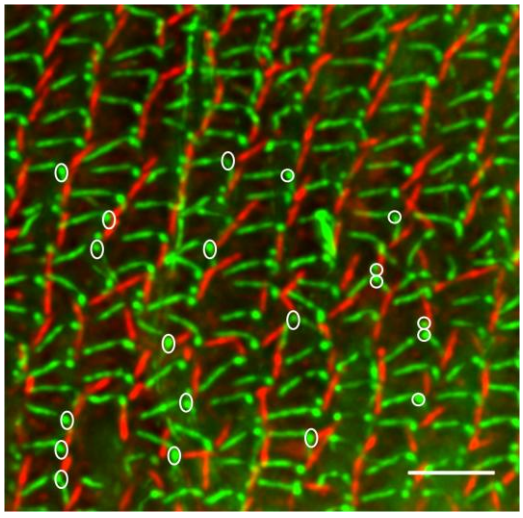
Panels show the staining of all three rootlets (TR (green), PR (green) and SR (red)) along with basal bodies (green). Panel A (control), panel B (Paralog Group: SR 2 depleted cell), panel C (Structural Group 1: Paralog Group SR 1) and panel D (Structural Group 2: Paralog Group SR 7) are showing the basal body and its associated rootlets in different cells. Scale bars are 3μm. White circles denote the position of the basal bodies that were used in measuring the angles between rootlets to show that basal bodies from disrupted rows were clearly represented in the angle measurements.



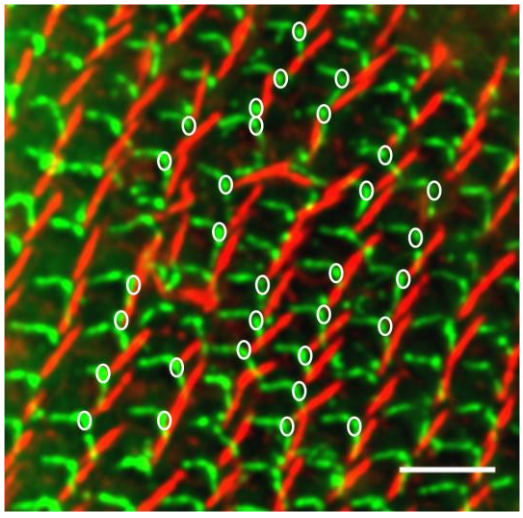
A (Control cell)



B (SR 2 depleted cell)



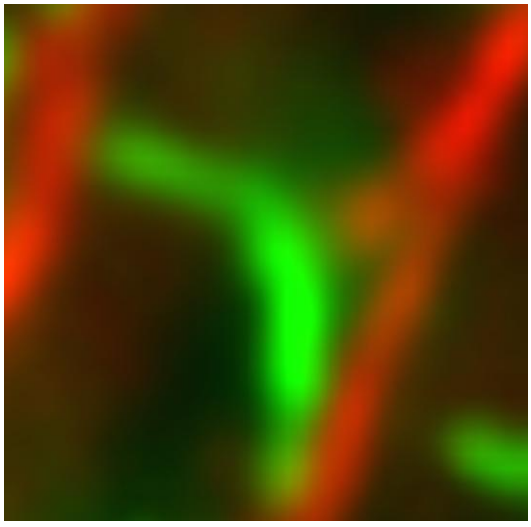
C (SR 1 depleted cell)



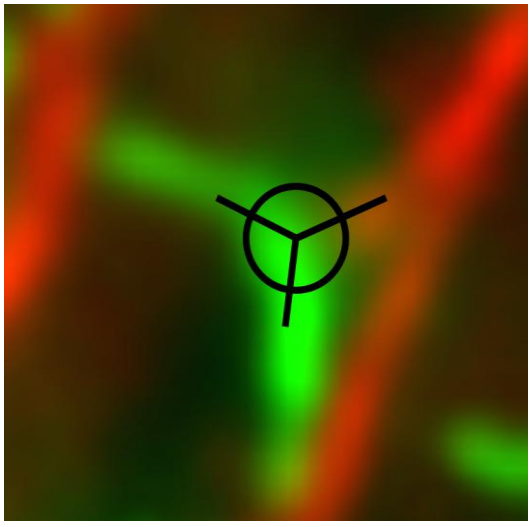
D (SR 7 depleted cell)

Fig. 2.11: PCR vs TR and PCR vs SR angles remain relatively constant in the control, SR Paralog Group or SR Structural Group depleted cells.

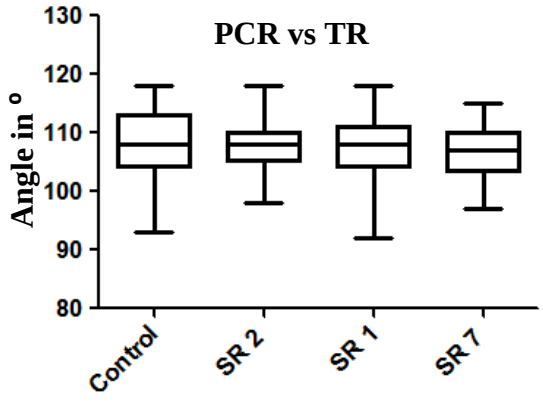
Panel A shows the enlarged image of an individual basal body with all three rootlets (TR (green), PR (green) and SR (red)) along with basal bodies (green). Panel B demonstrates how the angles were measured. First a circle was drawn around the basal body; then three straight lines were drawn to determine the positions of rootlets to the basal body. Finally, angles were measured by protractor. Image C and D show the measurement of angles between PR vs TR and PR vs SR in control, Paralog Group SR 2, Structural Group1 (Paralog Group: SR 1) or Structural Group 2 (Paralog Group: SR 7) depleted cells respectively.



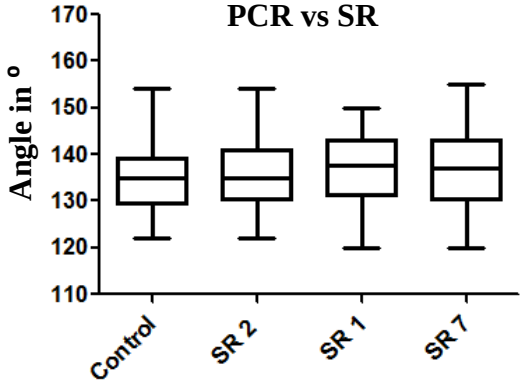
A



B


PCR vs TR
Angle in °
130
120
110
100
90
80
Control
SR 2
SR 1
SR 7

C

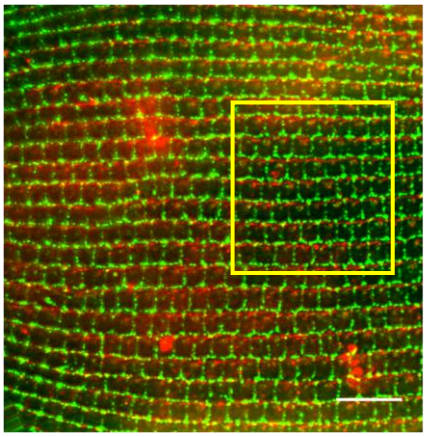

PCR vs SR
Angle in °
170
160
150
140
130
120
110
Control
SR 2
SR 1
SR 7

D

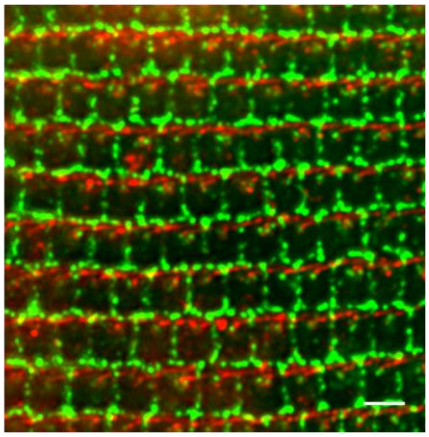
Fig.2.12: SR Structural Group depletion causes distorted cortical units with abnormal SRs.

Panels A, C and E show control, SR Paralog Group SR 2 and SR Structural Group 2 (Paralog Group: SR 7) depleted cells respectively. The yellow box in each image is enlarged (B, D and E) to show basal body rows (green) with SRs (red). All images are of the dorsal surface. Panels B and D show organized cortical units as well as SRs that extend from the posterior pole towards the anterior in the control and SR Paralog Group SR 2 depleted cells respectively. Panel F shows the severely distorted cortical units and SRs with abnormal appearances. Distortion of cortical units with abnormal SRs occurs all over the entire cell surface. Scale bars are 10μm (A, C and E) and 3μm (B, D and F).

Control cell

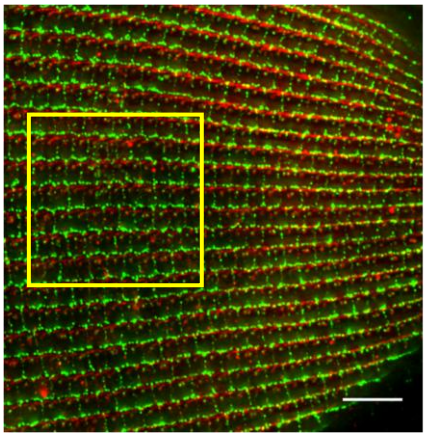


A

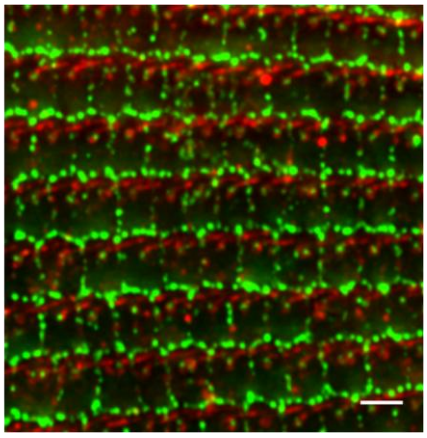


B

SR 2 depleted cell

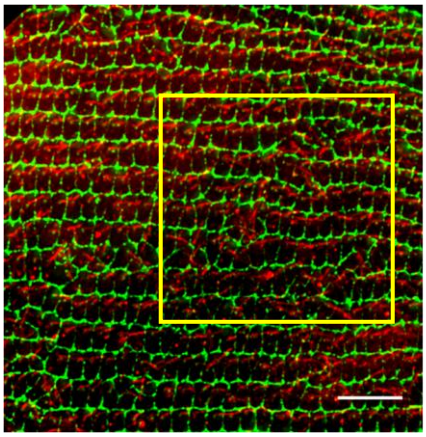


C

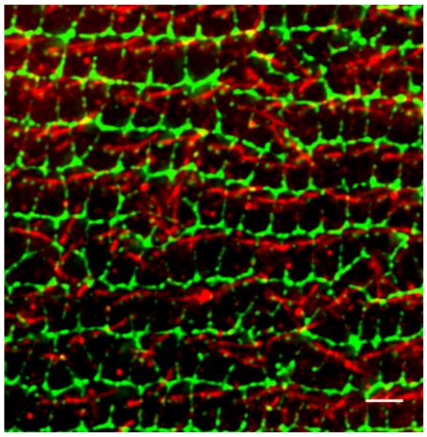


D

SR 7 depleted cell



E

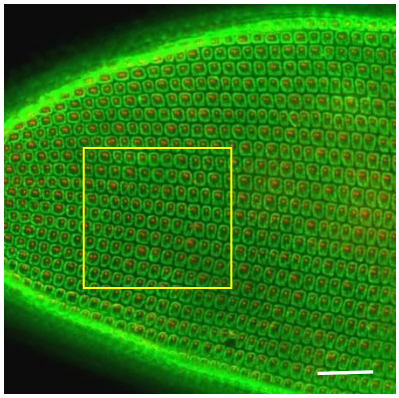


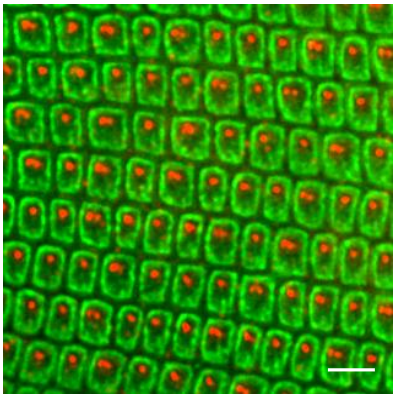
F

Fig.2.13: Depletion of SR Structural Group causes irregular shaped epiplasm with misaligned basal body rows.

Panel A, C and E show control, SR Paralog Group SR 2 and SR Structural Group 2 (Paralog Group: SR 7) depleted cells respectively. The yellow box in each image is enlarged (B, D and E) to show epiplasm territories with basal body units. The epiplasm territories are green (anti-epiplasm) and basal body units are red (anti-centrin). All images are of the dorsal surface. Panel B and D show regularly shaped territories of epiplasm with basal body unit at the center of the each epiplasm territory in the control and SR Paralog Group SR 2 depleted cells. Panel F shows the severely misshaped epiplasm territories with misaligned basal body rows. Scale bars are 10μm (A, C and E) and 3μm (B, D and F).

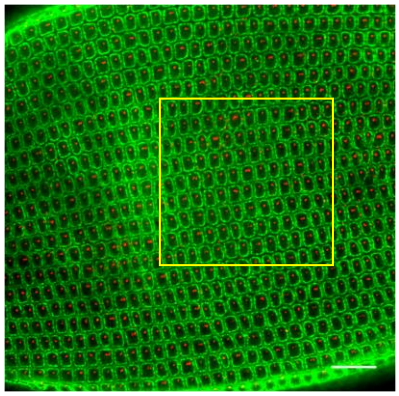
Control cell

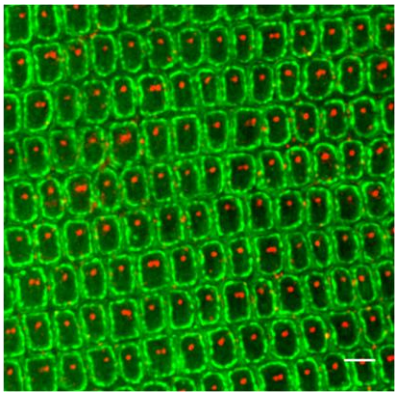




A **B**

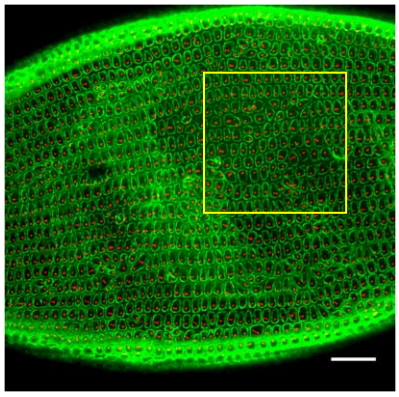
SR 2 depleted cell

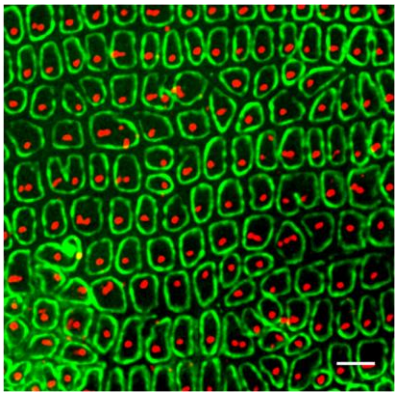




C **D**

SR 7 depleted cell



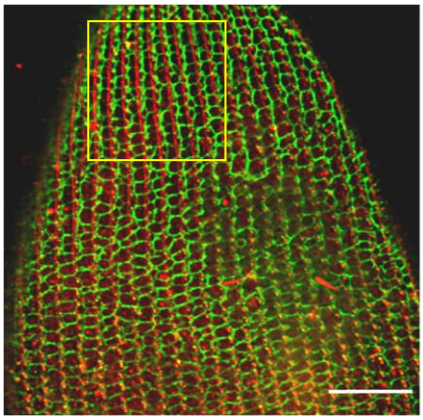


E **F**

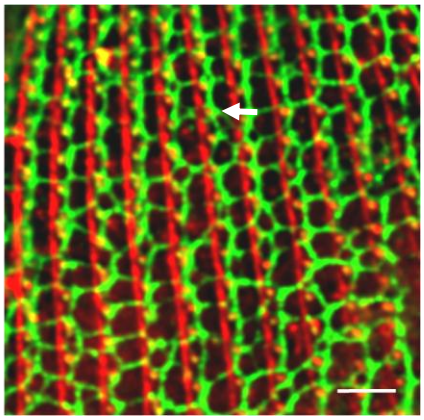
Fig.2.14: Depletion of SR Structural Group affects the association between ICL network and SRs along with misaligned basal body rows.

Panel A, C, and E show control, Paralog Group SR 2 and Structural Group 2 (Paralog Group SR 7) depleted cells respectively. All of these cells were expressing FLAG-Ptcen15. Anti-Flag staining of these cells decorate ICL network (green). Anti-centrin and anti-SR decorate the basal body units (red) and SRs (red) respectively in these cells. Panel E and G show the basal body units (red) along with SRs (red) staining in FLAG-empty vector expressing cells. In these cells there was so signal for FLAG epitope. The yellow box in each image is enlarged (B, D and E) to ICL network along with basal body units (red) and SRs (red). All images are of the dorsal surface. Panel B and D show close association of ICL network and basal body units along with SRs (white arrows). Panel F shows the association is somehow missing in between the ICL network and basal body units with abnormal SRs (white arrows). Scale bars are 10μm (A, C, E, G and I) and 3μm (B, D, F, H and J). All images are flipped to analyze the basal body with its' associated SR and ICL organization easily. Please note that the focal plane of the basal body units, SRs and ICL network in the images are very different respective to each other.

Control cell expressing Flag-Ptcen15

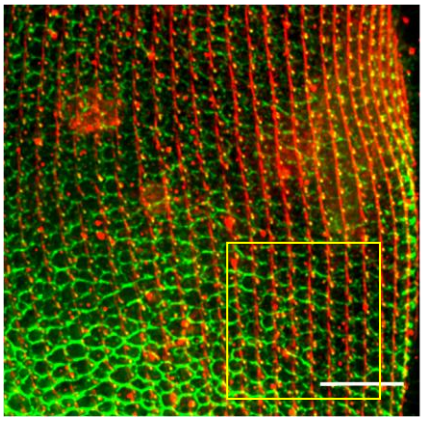


A

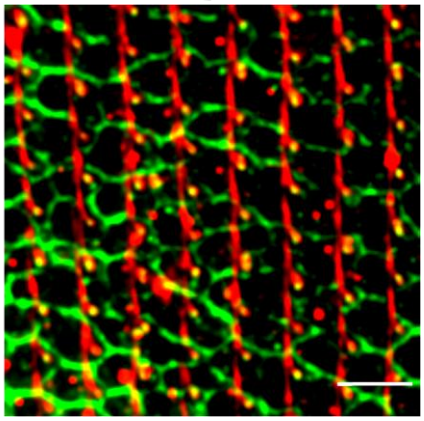


B

SR 2 depleted cell expressing Flag-Ptcen15

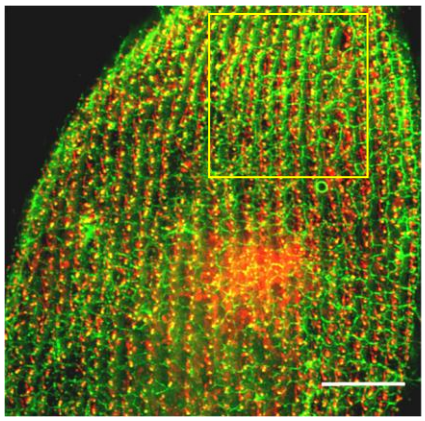


C

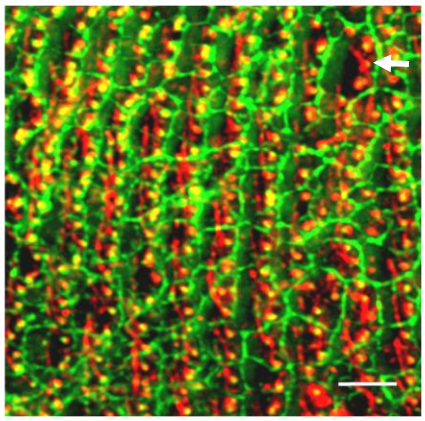


D

SR 7 depleted cell expressing Flag-Ptcen15

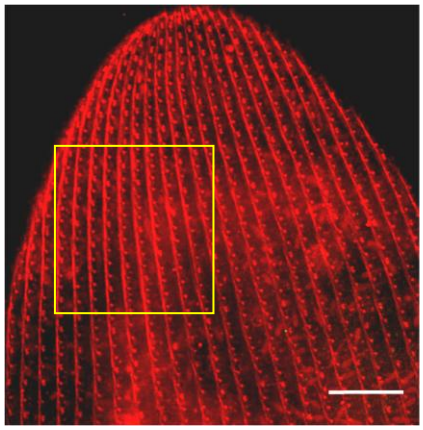


E

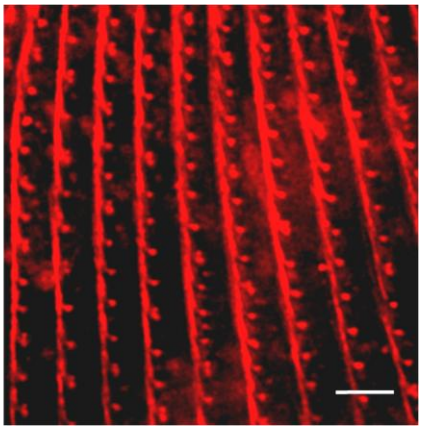


F

Control cell expressing Flag-empty vector

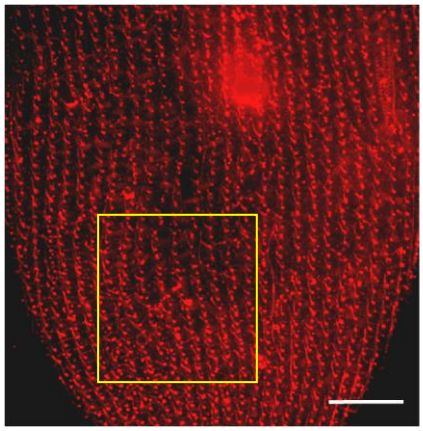


G

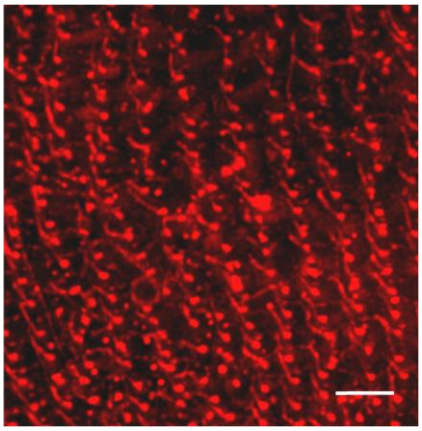


H

SR7 depleted cell expressing Flag-empty vector



I

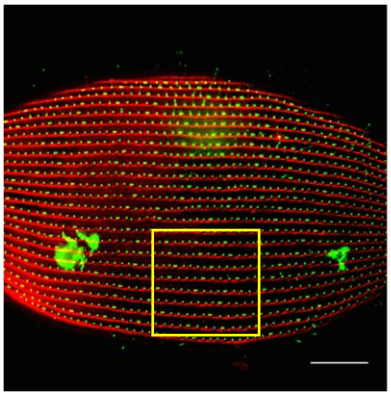


J

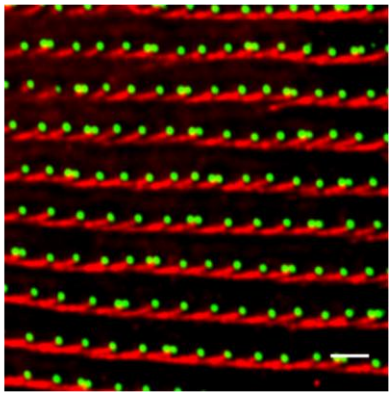
Fig.2.15: Depletion of ICL 1e subfamily proteins causes misaligned basal body rows with a disordered organization of the SRs.

Panel A, C and E show ICL9, ICL10 and ICL1e depleted cells respectively. The yellow box in each image is enlarged (B, D and E) to show basal body rows with SRs. The basal bodies are green (anti-ID5) and SRs are red (anti-SR). All images are of the dorsal surface. Panel B and D show straight rows of basal bodies as well as SRs that extend from the posterior pole towards the anterior. Panel F shows the severely misaligned basal body rows and the highly disordered SRs. Misalignment occurs over the entire cell surface. Scale bars are 15μm (A, C and E) and 3μm (B, D and F).

ICL 9 depleted cell

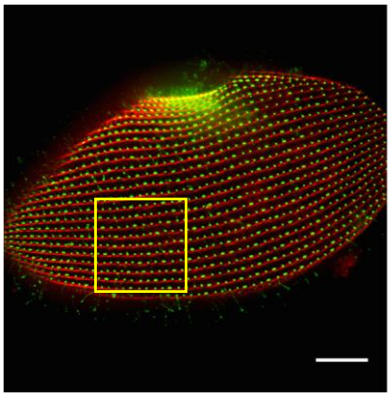


A

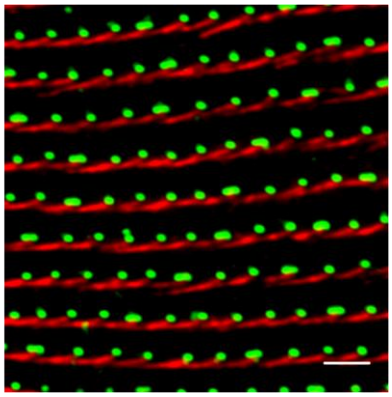


B

ICL 10 depleted cell

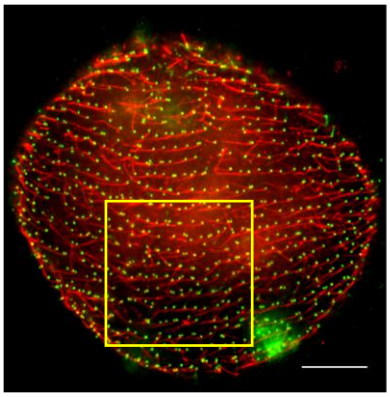


C

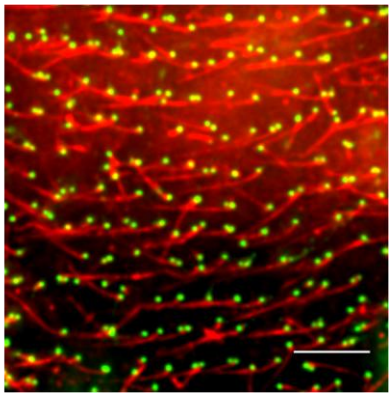


D

ICL 1e depleted cell



E



F

Table 2.1: RNAi constructs and nucleotides

Name	Number	Nucleotide base position in gene sequence
SR 1a	GSPATG00025723001	+253 - +842
SR 1b	GSPATG00023935001	+257 - +870
SR 2	GSPATG00020957001	+23 - +790
SR 3	GSPATG00022867001	+163 - +786
SR 4	GSPATG00008626001	+3 - +805
SR 5a	GSPATG00026454001	+27 - +632
SR 5b	GSPATG00024123001	+7 - +716
SR 6a	GSPATG00010780001	+341 - +800
SR 6b	GSPATG00014274001	+114 - +726
SR 6d	GSPATG00014806001	+174 - +632
SR 7a	GSPATG00036966001	+178 - +680
SR 8a	GSPATT00032447001	+14 - +864
SR 8b	GSPATG00033989001	+64 - +740
SR 9	GSPATT00011977001	+59 - +862
SR 10a	GSPATG00034929001	+113 - +439
SR 10d	PTETG17800004001	+364- +820
SR 11a	GSPATG00002032001	+39 - +690
SR 11b	GSPATG00017545001	+115 - +698
SR 12a	GSPATG00001334001	+132 - +793
SR 12b	GSPATG00002707001	+87 - +791
SR 12d	GSPATG00004838001	+141 - +728
SR 13a	GSPATG00025686001	+127 - +799
SR 13c	GSPATG00020721001	+122 - +712
SR 13d	GSPATG00021965001	+116 - +694
PtCen_icl1e	GSPATP00007681001	+13 - 598
PtCen_icl9a	GSPATP00027865001	+1 - +679
PtCen_icl10a	GSPATP00003368001	+1 - +643

Table 2.2: Primers used in the study

Name		Primers 5'to 3'
SR 1a	Forward	GGCAGACGACTCAAAATTTAAG
	Reverse	GACTCTCTTCAGTTGCTTCTC
SR 1b	Forward	GGAGATGATTCTAAGTTCAAG
	Reverse	CATGTTTGATCGATCAAATTG
SR 2	Forward	GAATAAGAACAGGAGAAG
	Reverse	CTTCTTGCATGCCTTTCC
SR 3	Forward	CATTAGTCAGATCAGAAGG
	Reverse	CACTAATGTTTCCTCTGTC
SR 4	Forward	GTCTGAATCCAATGTTAAAG
	Reverse	GTACAAGAGTCTCTTCTGTC
SR 5a	Forward	GACAATGTATAAACTAGAA
	Reverse	CAAATAATTCTCTTAGGTCTG
SR 5b	Forward	GACAATCTATAAACTAGAGTA AC
	Reverse	CAGTAATTCTAAATTACACTTCTC
SR 6a	Forward	GATTGATCGAATAGAATTCC
	Reverse	CGACTAATTAGACTAAGG
SR 6b	Forward	CGATTCCATGCTATCAGCC
	Reverse	GATTTCTCTCAGACTTTCCAC
SR 6c	Forward	CATCTCTACTAGTGTTTAGG
	Reverse	CCTTCTCAACCAACTGCAC
SR 6d	Forward	GAGGCTTAATTGTAACATC
	Reverse	CGACAAATTAATGGTATATC
SR 7a	Forward	GAGGAATCAGAGACAAGATTAG
	Reverse	GTCTTTATAGCTTCATGTAATC
SR 8a	Forward	GATCCTAAGTCTCATGTATC
	Reverse	CTGCAATACTGAGTTTATTAC
SR 8b	Forward	GCATCACCATCAAGTGGAAG
	Reverse	CTCTATTGACGATATCTTTAAGC
SR 9	Forward	GTCTCCAAGCAGTGGAAG
	Reverse	CTCTGTTTCTCTTCTCCAC
SR 10a	Forward	CCAGATAGGAGTGGAGAATG
	Reverse	CCTCTGTCAATCTAAGAATC
SR 10b	Forward	GATCTCAGAGAAGCTGTCAAC
	Reverse	CTGGTTCTCTTCTCTTGATCC
SR 10c	Forward	CTGCAGTGTAATATTGTGAGTG
	Reverse	CAGATTGCTCTGTGGTTCTC
SR 10d	Forward	GACCAAGAAGTTCGAGAATGC
	Reverse	CTAGTGTAATGAAGATGGATACC
SR 11a	Forward	CAATCCAACTACATCACGATAG
	Reverse	GCATCTCAAATATTGCTGATTCAC
SR 11b	Forward	CGTGTTTCTGGTGTAGAAGAG

	Reverse	CCTTGAGCATCTCAAATATTGC
SR 12a	Forward	GCGGCTAGCGGAGAGAATTTAAGACAC
	Reverse	GCGGGTACCCAAGCATCTTCCAATAAT
SR 12b	Forward	GGCACACAGAATAGATACGG
	Reverse	CATGCATCTTCTAACAGTCC
SR 12c	Forward	GAGAGAATGCAAGACACG
	Reverse	CAAGCATCT TCTAGAAGTCC
SR 12d	Forward	GCAAGACACAATTACCTC
	Reverse	CGATTTCGCTCTTGACTC
SR 13a	Forward	GTAGAGGAGAGAACTCAAG
	Reverse	CAGCATCTTCGAGTAATCC
SR 13c	Forward	GCAGTAGAGGAGAGAATTCAG
	Reverse	CAAAGATAGCCGTCTCGG
SR 13b	Forward	CAGCAGTTGAAGAGAGAAC
	Reverse	CTAGATCCTTGAGCATCTC
SR 13d	Forward	GAGTCACAGCAGTAGAAGAGAG
	Reverse	CATTTCAAAGATTGCTGTCTCC
SR 1a Flag	Forward	CGCGGGCCCATGTCTGCATCTAAAATC
	Reverse	CGCGAGCTC TCATTTTCTTAATGAATTTTC
SR 2 Flag	Forward	GCGGCTAGCATGTCTGAGAATATCTAAAATAGAG
	Reverse	GCGGGTACCTCATAAATTCATTTCTAACATTTCTC
SR 3 Flag	Forward	GCGGCTAGCATGTCTGAAAGTATCTAAAATAG
	Reverse	GCGGGTACCTCATAAATTCATTTCTAACATTTC
SR 4 Flag	Forward	CGCGGGCCCATGTCTGAATCCAATGTTA
	Reverse	CGCGAGCTCTCATAAATTCATTTCTAAC
SR 5a Flag	Forward	CGCGGGCCCATGAATGACAATCTATAAAC
	Reverse	CGCGAGCTCTCACATATTCATTTCTAAC
SR 6b Flag	Forward	GCGGGTACCTCATAAGTTCATTTCTAACATTTC
	Reverse	GCGGAGCTCATGTCAGAAAATATCTAAAATAGAG
SR 7a Flag	Forward	GCGGCTAGCATGTCAAATAATAGAAG
	Reverse	GCGGGTACCTCATGCTAGTTAAGC
SR 8a Flag	Forward	GCGGCTAGCATGTACAATTCAAGATCCTAAG
	Reverse	GCGGGTACCTCATAAATTTGCTGCAATACTG
SR 9 Flag	Forward	GCGGCTAGCATGTATAATTCAAGATC
	Reverse	GCGGGTACCTCATAGATTTGCTGC
SR 10a Flag	Forward	GCGGCTAGCATGTCCAGAGTCAAACAATC
	Reverse	GCGGGTACCTCATAAATTTGCTGCAATGC
SR 11a Flag	Forward	GCGGCTAGCATGAATCCAAATACACCATTCAACG
	Reverse	GCGGGTACCTCAATCTTTTGCTGCTCTGTAAATC
SR 12a Flag	Forward	GCGGCTAGCATGAATCCAAATACACCCTTC
	Reverse	GCGGGTACCTCAGTCTTTTGCTGCTCTG
SR 13a Flag	Forward	GCGGCTAGCATGAATCCCAATACACCTTTC
	Reverse	GCGGGTACCTCAATCCTTGGCTGCTCTG
Ptcen_icl1e	Forward	ATGAGCAAAAAACAATAAGC

	Reverse	TCAGAATGTTCTTTTGACC
Ptcen_icl9a	Forward	ATGAATAGAAATCAATCCAG
	Reverse	TCATGCTTGATTTGGGTAT
Ptcen_icl10a	Forward	ATGAATAGAAAGAAAAATAC
	Reverse	TCAATTTTTCTTTTGTAG
Ptcen15 Flag	Forward	CGCGGGCCCATGAGAGGTAAGCAATAG
	Reverse	CGCGAGCTCTCAAAATGTTCTTTTAACC
Calmodulin	Forward	CTGAAGCTGAACTTCAAG
	Reverse	CAGAATGATGGTTTCTAAATGA

Table 2.3: Position of the coiled coil domains in the SR proteins

Name	Number	AA	MW kDa	CC1	CC1 #AA	CC2	CC2 #AA
SR 1a	GSPATP00025723001	272	31.81	30-58	29	161-226	66
SR 1b	GSPATP00023935001	274	31.99	32-60	29	163-228	66
SR 7a	GSPATP00036966001	251	29.27	31-63	33	175-240	66
SR 7b	GSPATP00034214001	251	29.27	31-63	33	175-240	66
SR 8a	GSPATP00032447001	257	29.72	34-119	86	213-243	31
SR 8b	GSPAT00033989001	257	29.70	34-119	86	213-243	31
SR 9	GSPATT00011977001	255	29.55	32-117	86	211-241	31
SR 10a	GSPATP00034929001	256	29.64	33-118	86	212-242	31
SR 10b	GSPATP00028132001	256	29.61	33-118	86	212-242	31
SR 10c	PTETP6800004001	256	29.57	33-118	86	212-242	31
SR 10d	PTETP17800004001	256	29.48	33-118	86	212-242	31
SR 11a	GSPATP00002032001	248	28.81	44-91	48	203-236	34
SR 11b	GSPATP00017545001	248	28.81	44-91	48	203-236	34
SR 12a	GSPATP00001334001	248	28.80	44-91	48	203-236	34
SR 12b	GSPATP00002707001	248	28.80	44-91	48	203-236	34
SR 12c	GSPATP00003161001	248	28.80	44-91	48	203-236	34
SR 12d	GSPATP00004838001	248	28.74	44-91	48	203-236	34
SR 13a	GSPATP00025686001	248	28.94	44-91	48	203-236	34
SR 13b	GSPATP00030915001	248	28.98	44-91	48	203-236	34
SR 13c	GSPATP00020721001	248	28.99	44-91	48	203-236	34
SR 13d	GSPATP00021965001	248	29.15	44-91	48	203-236	34
SR 6a	GSPATP00010780001	249	29.89	174-232	59		
SR 6b	GSPATP00014274001	249	29.91	174-232	59		
SR 6c	GSPATP00005889001	249	29.86	174-232	59		
SR 6d	GSPATP00014806001	249	29.99	174-232	59		
SR 2	GSPATP00020957001	249	29.82	174-232	59		
SR 3	GSPATP00022867001	249	29.89	174-232	59		
SR 4	GSPATP00008626001	258	30.57	183-241	59		
SR 5a	GSPATP00026454001	249	29.93	174-232	59		
SR 5b	GSPATP00024123001	249	30.04	174-232	59		

CC1 – Coiled-coil domain 1

CC2 – Coiled-coil domain 2

#AA – Number of amino acids spanning the coiled-coil domain

Table 2.4: SR proteins found by LC-MS/MS among proteins in the 30% Optiprep fraction in silver stained gels in the mass range (~38-29 kDa)

	SR protein	Total peptide sequences identified	Unique to the paralog
SR 1 (Structural Group 1)	SR 1a	28	8
	SR 1b		13
SR 2-6 (Structural Group 5)	SR 2	60	5
	SR 3		1
	SR 4		5
	SR 5a		2
	SR 5b		4
	SR 6a		0
	SR 6b		0
	SR 6c		3
	SR 6d		3
SR 7 (Structural Group 2)	SR 7a	13	0
	SR 7b		0
SR 8-10 (Structural Group 3)	SR 8a	35	1
	SR 8b		0
	SR 9		3
	SR 10a		0
	SR 10b		3
	SR 10c		0
	SR 10d		2
SR 11-13 (Structural Group 4)	SR 11a	47	1
	SR 11b		2
	SR 12a		0
	SR 12b		0
	SR 12c		0
	SR 12d		0
	SR 13a		2
	SR 13b		1
	SR 13c		6
	SR 13d		0

Table 2.5: Summary table

	RNAi Phenotype	SR striation	SR length
SR 1	BB row misalignment	Absent	Short
SR 2	BB row normal	Normal	Normal
SR 3	BB row normal	Normal	Normal
SR 4	BB row normal	Normal	Normal
SR 5	BB row normal	Normal	Normal
SR 6	BB row normal	Normal	Normal
SR 7	BB row misalignment	Absent	Short
SR 8	BB row normal	Normal	Normal
SR 9	BB row normal	Normal	Normal
SR 10	BB row normal	Normal	Normal
SR 11	BB row normal	Normal	Normal
SR 12	BB row normal	Normal	Normal
SR 13	BB row normal	Normal	Normal
SR 2-6	BB row misalignment	Absent	Short
SR 8-10	BB row misalignment	Absent	Short
SR 11-13	BB row misalignment	Absent	Short

Supplement Data

2.S.1: Amino acid sequences alignment of coiled coil domains of different Structural Groups .

Panels A, B, C, D and E show the amino acid sequences alignment of coiled coil domains of Structural Group 1, Structural Group 2, Structural Group 3, Structural Group 4 and Structural Group 5 respectively.

A

Alignment of coiled coil domain 1 in Structural Group 1

```
SR1a    30    SRAKISQLSEKLSNLQHSIDEDQAFKKET
SR1b    32    SRAKISQLSEKLSNLQHQIDEDQAFKKET
              ****************.***********
```

Alignment of coiled coil domain 2 in Structural Group 1

```
SR1a    161   REETEEKNAQEIGDRVLQLQEEVEEERRQREQQNQDTIKRLGDSILKLQEILTTEKKQREQAQSQMQAQSQM
SR1b    163   REETEEKNAQEIGDRILQLQEEVEEERRQREQQNQNTIKRLGDSILKLQEILTTEKKQREQAQGQMQAQGQM
              ***************:******************:*************************.*****.**
```

B

Alignment of coiled coil domain 1 in Structural Group 2

```
SR7a    31    ERTKRIEESETRLDQLESGLNELNEQLSQRVAL
SR7b    31    ERTKRIEESETRLDQLESGLNELNEQLSQRVAL
              ********************************
```

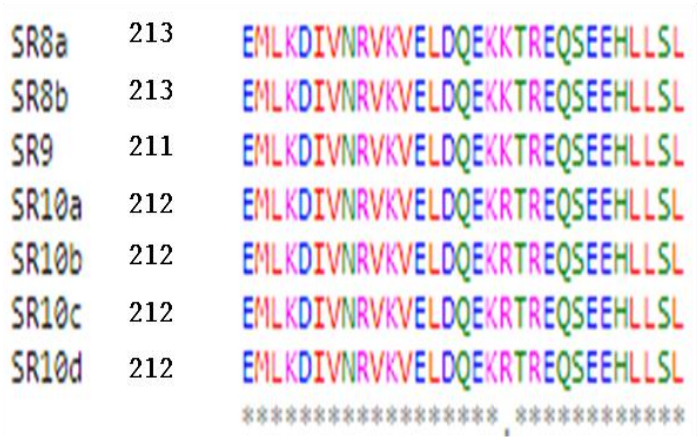
Alignment of coiled coil domain 2 in Structural Group 2

```
SR7a    175   NISRKMNEELSKLNQLLVQEKVNRQESEQAIFDMLKDVVNRIKTEVDNEKKQRESTEETLLALLED
SR7b    175   NISRKMNEELSKLNQLLVQEKVNRQESEQAIFDMLKDVVNRIKTEVDNEKKQRESTEETLLALLED
              ************************************************************************
```

C

Alignment of coiled coil domain 1 in Structural Group 3

Alignment of coiled coil domain 2 in Structural Group 3



D

Alignment of coiled coil domain 1 in Structural Group 4

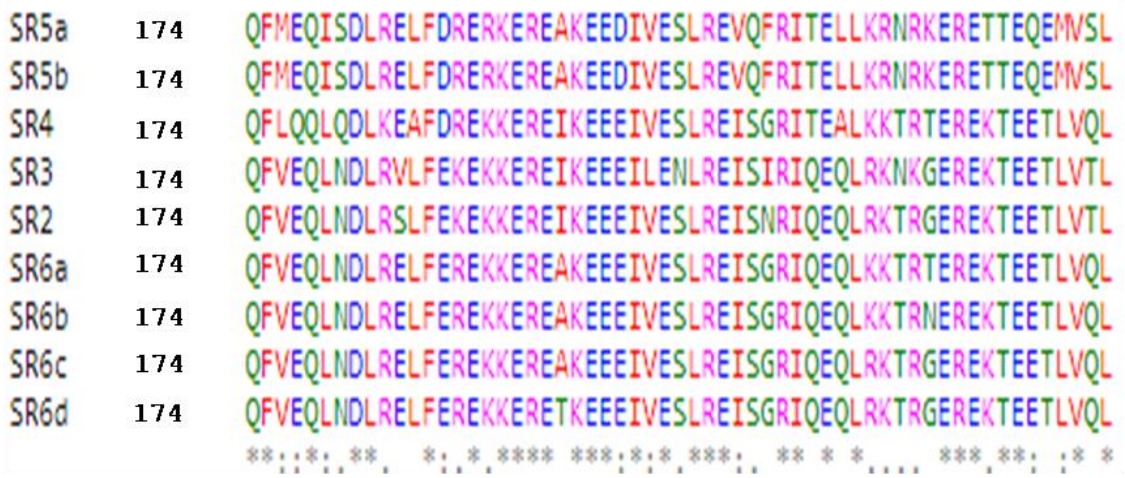
```
SR13d   44   EERIQDTITSYNRKLHTLKDEIVRVQKQIEEENNTFETQFEQRVREIA
SR12c   44   EERMQDTITSYNRKLHALKDEIVRLQKQIEEENNAFETQFEQRVREIA
SR13b   44   EERTQDTITSYNRKLHSLKDEIVRFQKQIEEENNAFETQFEQRVREIA
SR13c   44   EERIQDTITSYSRKLHSLKDEIVRLQKQIEEENNAFETQFEQRVREIA
SR13a   44   EERTQDTITSYNRKLHSLKDEIVRLQKQIEEENNAFETQFEQRVREVA
SR12d   44   EERMQDTITSYNRKLHSLKDEIVRLQKQIEEENNAFETQFEQRVREVA
SR11a   44   EERIQDTITSYNRKLHSLKDEIVRLQKQIEEENNAFETQFEQRVREVA
SR11b   44   EERIQDTITSYNRKLHSLKDEIVRLQKQIEEENNAFETQFEQRVREVA
SR12a   44   EERIQDTITSYNRKLHTLKDEIVRLQKQIEEENNAFETQFEQRVREVA
SR12b   44   EERIQDTITSYNRKLHTLKDEIVRLQKQIEEENNAFETQFEQRVREVA
             *** ******* **** ******* ******** ********** *
```

Alignment of coiled coil domain 2 in Structural Group 4

```
SR12a   203   FEMLKDLVSRVKSEIEEEKKLREESQESLLGLLE
SR12b   203   FEMLKDLVSRVKSEIEEEKKLREESQESLLGLLE
SR12c   203   FEMLKDLVSRVKSEIEEEKKLREESQESLLGLLE
SR12d   203   FEMLKDLVSRVKSEIEEEKKLREESQESLLGLLE
SR13a   203   FEMLKDLVSRVKSEIEEEKKLREESQESLLGLLE
SR13b   203   FEMLKDLVSRVKSEIEEEKKLREESQESLLGLLE
SR13c   203   FEMLKDLVSRVKSEIEEEKKLREESQESLLGLLE
SR13d   203   FEMLKDLVSRVKSEIEEERKLREESQESLLGLLE
SR11a   203   FEMLKDLVSRVKSEIEEEKKLREESQENLLGLLE
SR11b   203   FEMLKDLVSRVKSEIEEEKKLREESQENLLGLLE
              ****************** ******** ******
```

E

Alignment of coiled coil domain 1 in Structural Group 5



2.S2: Amino acid sequences alignment of 30 SR proteins.

Alignment of all 30 SR proteins in *Paramecium*:

```
SR5a            IIAQLIGQFESNLRNASQQKEAVFSQLANKFKELQVFLESENDNRVRQEQE-TQKMIVDL
SR5b            IIAQLIAQFESNLRNASQQKEAIFSQLANKFKELQVFLEGENDNRVRQEQE-TQKMIVDL
SR4             NISQLIMAFEAQLQHNSEQKEEKFNYLASRMRQLGEFLEQEQEDRLRQEAE-TQKLLTDL
SR6d            TISQLIMAFEAQLQHQSDQKEEKFAYIAQKVRQITEFLEQEQEDRERQESE-TFKLITDL
SR6c            AISQLIMAFEAQLQHQSDLKEEKFAYIAQKVRQITEFLEQEQEDRERQESE-TFKLITDL
SR6a            AISQLIMAFEAQLQHQSDQKEEKFAYIAQKVRQLTEFLEQEQEDRERQESE-TFKLITDL
SR6d            AISQLIMAFEAQLQHQSDQKEEKFAYIAQKVRQLTEFLEQEQEDRERQESE-TFKLITDL
SR3             AISQLIMAFEAQLQHQSDQKEEKFIYIAAKVRQITEFLEQEQENRERQETE-TFKLITDL
SR2             AISQLIMAFEAQLQHQSDQKEEKFTYIAAKVRQITEFLEQEQENRERQETE-TFKLITDL
SR1a            TFESKVKILEDKATKQAQADDSKFKLLKEQLQKVEEGAQNEKIIREAGDEKLRSKDLKSL
SR1b            TFESKIKILEDRATKQAQGDDSKFKLLKEQLQKVEEGAQNEKIIREAGDEKLRSKDLKGL
SR13a           TICGRITAVEERTQDTITSYNRKLHSLKDEIVRLQKQIEEENNAFETQFEQ-RVREVAAF
SR13b           TICGRITAVEERTQDTITSYNRKLHSLKDEIVRFQKQIEEENNAFETQFEQ-RVREIAAF
SR11a           VICGRVSGVEERIQDTITSYNRKLHSLKDEIVRLQKQIEEENNAFETQFEQ-RVREVAAF
SR11b           VVCGRVSGVEERIQDTITSYNRKLHSLKDEIVRLQKQIEEENNAFETQFEQ-RVREVAAF
SR12a           TVCGRITGVEERIQDTITSYNRKLHTLKDEIVRLQKQIEEENNAFETQFEQ-RVREVAAF
SR12b           TVCGRITGVEERIQDTITSYNRKLHTLKDEIVRLQKQIEEENNAFETQFEQ-RVREVAAF
SR12c           TICGRITGVEERMQDTITSYNRKLHALKDEIVRLQKQIEEENNAFETQFEQ-RVREIAAF
SRG12d          NICGRITGVEERMQDTITSYNRKLHSLKDEIVRLQKQIEEENNAFETQFEQ-RVREVAAF
SR13d           TINSRVTAVEERIQDTITSYNRKLHTLKDEIVRVQKQIEEENNTFETQFEQ-RVREIAAF
SR13c           TINSRVTAVEERIQDTITSYSRKLHSLKDEIVRLQKQIEEENNAFETQFEQ-RVREIAAF
SR7a            ESETRLDQLESGLNELNEQLSQRVALIKDSVLKLQKVLDQTKLQREQHFEQ-KQKEYIDL
SR7b            ESETRLDQLESGLNELNEQLSQRVALIKDSVLKLQKVLDQTKLQREQHFEQ-KQKEYIDL
SR9             QAEQRIQQAEDAFNEFQEQIFTKLNGLRDQLGKLQKQVEEDKLAKEQANET-KNREVLAL
SR8a            QAEQRIQQAEDAFNEFQEQIFTKLNGLRDQLGKLQKQVEEDKLAKEQANET-KNREVQAL
SR8b            QAEQRIQQAEDAFNEFQEQIFTKLNGLRDQLGKLQKQVEEDKLAKEQANET-KNREVQAL
SR10a           QAEQRIQQAEDAFNEFQEQIFTRLNGLRDQLGKLQKQVEDDKQAKEQAIEA-KNRDVLAL
SR10b           QAEQRIQQAEDAFNEFQEQIFTRLNGLRDQLGKLQKQVEDDRQAKEQAIEA-KNRDVQAL
SR10c           QAEQRIQQAEDAFNEFQEQIFTKLNGLRDQLGKLSKQVEDDRLAKEQANEA-KNRDVQAL
SR10d           QAEQRIQQAEDAFNEFQEQIFTKLNGLRDQLGKLSKQVEDDRLAKEQANEA-KNREVQAL
                     :  *            . :  . .  : .           .   :
```

```
SR5a            ERHAKRLLETSQKERVDQEKKLTYAISQQIDFISQDVNRQGIE--------LLESHKYVD
SR5b            ERHAKRLLETSQKERVDQEKKLTYAISQQIDYIYQDVNRQGVE--------LLESHKYVD
SR4             ERHARKLIEQNSKDRVEQERKIVYQIGQQVDNLQGEVVKEGLVTKYNNIQAQSESYQFVD
SR6d            ERHARRLIEQNSKDRVEQEKKIVYSIGQQIENLQQDVIKEGLA--------QQTSHEYID
SR6c            ERHARRLIEQNSKDRVEQEKKIVYSIGHQIESLQQEVVKEGLA--------QSTSHEYID
SR6a            ERHARRLIEQNSKDRVEQEKKIVYTIGQQIESLQQEVVKEGLA--------QSTSHEYID
SR6d            ERHARRLIEQNSKDRVEQEKKIVYTIGQQIESLQQEVVKEGLA--------QSTSHEYID
SR3             ERHARRLIEQNSKERVEQEKKIVYSIGQQIESLQQDVIKEGLA--------QSTSHEYID
SR2             ERHARRLIEQNSKERVEQEKKVVYSIGQQIESLQQDVTKEGLA--------QSISHEYID
SR1a            EGFLNRELQQEKVNRKDFEGKIIKNTDDKVYSLRLDLARQKKY--------REETEEKNA
SR1b            EGFLNKELQSEKINRKDFEGKIIKNTDDKVYSLKLDLARQKKY--------REETEEKNA
SR13a           ESRITTKLEQEIALRRDGNLKLQGYLDEKVVYLKSDIQTEGKI--------RQEQTENIT
SR13b           ESRITTKLEQEIALRRDGNLKLQGYLDEKVVYLKSDIQTEGKI--------RQEQIENIT
SR11a           ESRITTKLEQEISIRKDGNLKLQGYLDEKVVYLKSDIQTESKI--------RQEQIENIT
SR11b           ESRITTKLEQEISIRKDGNLKLQGYLDEKVVYLKSDIQTESKI--------RQEQIENIT
SR12a           ESRITTKLEQEIALRKDGNLKLQGYLDEKVVYLKSDIQTEGKI--------RQEQVENIT
SR12b           ESRITTKLEQEIALRKDGNLKLQGYLDEKVVYLKSDIQTEGKI--------RQEQVENIT
SR12c           ESRITTKLEQEIALRKDGNLKLAGYLDEKVVYLKSDIQTEGKI--------RQEQIENIT
SRG12d          ESRITTKLEQEIALRKDGNLKLQGYLDEKVVYLKSDTQTEGKI--------RQEQIENIT
SR13d           ESRITTKLEQEIALRRDGNLKLQGYLDEKVVYLKSDIQTESKI--------RQEQIENIT
SR13c           ESRITTKLEQEIAIRRDGNLKLQGYLDEKVVYLKSDIQTEGKI--------RSEQIENIT
SR7a            ENSFNQAVEGLTSTRKDGEQKIIRFIEEKTGLIRSELSTESRT--------RNENIERLN
SR7b            ENSFNQAVEGLTSTRKDGEQKIIRFIEEKTGLIRSELSTESRT--------RNENIERLN
SR9             EKKFENAIENENQARKEGESKVLRLLEDKTALLRTEVQKETAS--------RVDAIEGIH
SR8a            EKKFENAIQNETQTRKEGETKVLRLLEDKTALLRTEVQKETAA--------RVDSIEGIH
SR8b            EKKFENAIQNETQTRKEGETKVLRLLEDKTALLRTEVQKETAA--------RVDAIEGIH
SR10a           TKKFENAIENEQQTKKEGEAKILRLTEDKSALLRTEVQKETAQ--------RIDAIEGIH
SR10b           TKKFENAIENEQQTKKEGEAKVLRLTEDKSALLRTEVQKETAQ--------RIDAIEGIH
SR10c           TKKFENAIENEQQTKKEGEAKVLRLTEDKTALLRTEVQKETAQ--------RVDAIEGIH
SR10d           TKKFENAIENEQQTKKEGEAKVLRLTEDKTALLRTEVQKETAQ--------RVDAIEGIH
                       ::     . : : *:      :   :  :   :               :
```

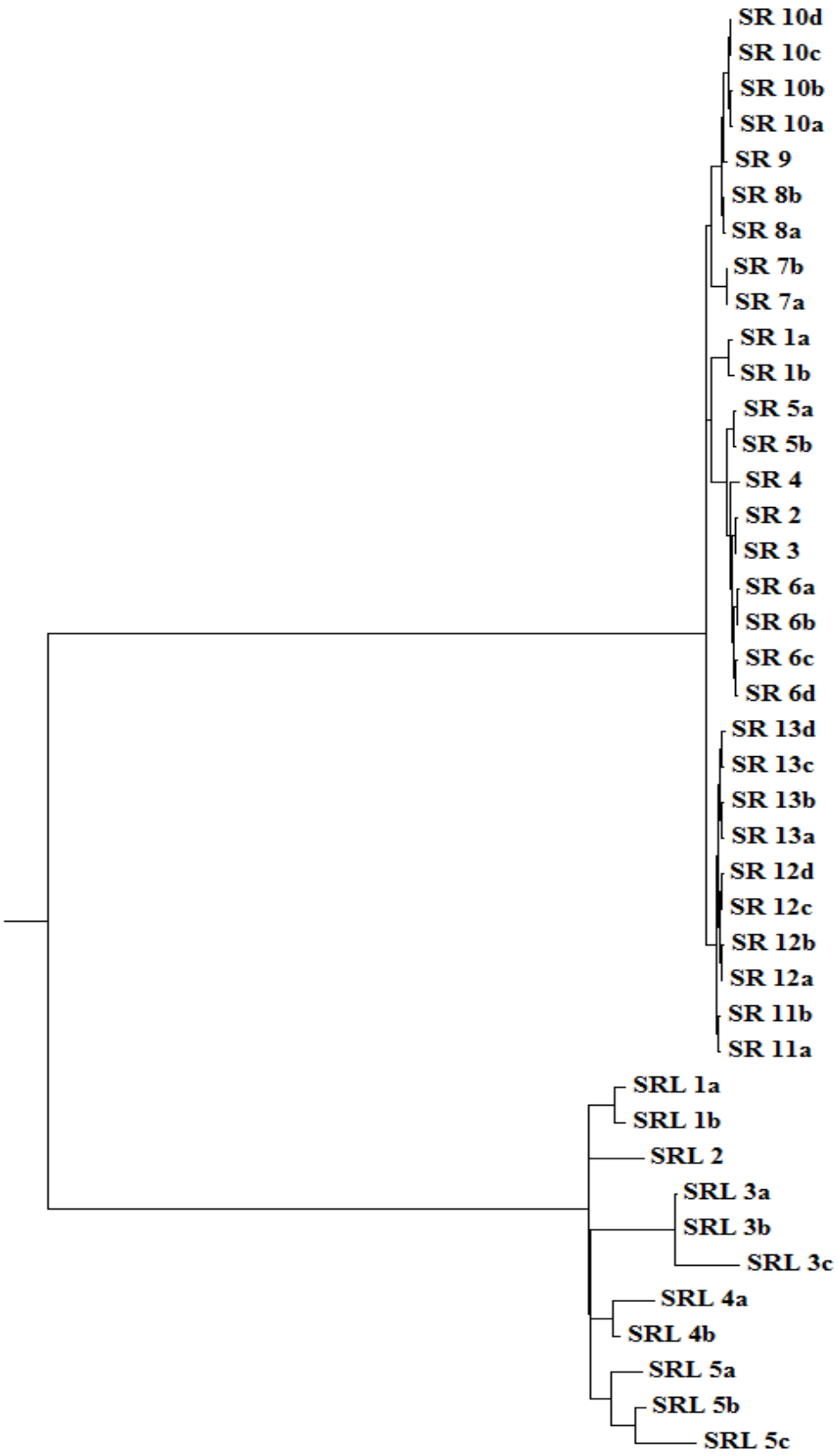
```
SR5a            IYLNEDLPKIADDLQNEIDERREVEERIYHQFMEQISDLRELFDRERKEREAKEEDIVES
SR5b            MYLNEDLPKIADDLQNEIDERREVEERIYHQFMEQISDLRELFDRERKEREAKEEDIVES
SR4             AYINEDLPKLADELQNEIGERKEVEGKIQIQFLQQLQDLKEAFDREKKEREIKEEEIVES
SR6d            SYLNEDLPKIADELQNEITERKDVEEKIYHQFVEQLNDLRELFEREKKERETKEEEIVES
SR6c            SYLNEDLPKIADELQNEITERKDVEEKIYHQFVEQLNDLRELFEREKKEREAKEEEIVES
SR6a            SYLNEDLPKIADELQNEITERKDVEEKIYHQFVEQLNDLRELFEREKKEREAKEEEIVES
SR6d            SYLNEDLPKIADELQNEITERKDVEEKIYHQFVEQLNDLRELFEREKKEREAKEEEIVES
SR3             SYLNEDLPKIADELQNEITERKDVEEKIYHQFVEQLNDLRVLFEKEKKEREIKEEEILEN
SR2             SYLNEDLPKIADELQNEITERKDVEENIYHQFVEQLNDLRSLFEKEKKEREIKEEEIVES
SR1a            QEIGDRVLQLQEEVEEERRQREQQNQDTIKRLGDSILKLQEILTTEKKQREQAQSQMFRM
SR1b            QEIGDRILQLQEEVEEERRQREQQNQNTIKRLGDSILKLQEILTTEKKQREQAQGQMFRM
SR13a           TSLENDLPKLYDMVKTEGQDREDSDNGTLRRAGDEIRRLNEGLGNQKKLREESETAIFEM
SR13b           TSLENDLPKLYDMVKTEGQDREDSDNGTLRRAGDEIRRLNEGVANQKKLREESETAIFEM
SR11a           TSLENDLPKLYDLVKTEGQDREDSDNGTLRRAGDEIKRLNEGLGNQKKLREESESAIFEM
SR11b           TSLENDLPKLYDLVKTEGQDREDSDNGTLRRAGDEIKRLNEGLSNQKKLREESEAAIFEM
SR12a           TSLENDLPKLYDMVKTEGQDREDNDNGTLRRAGDEIKRLNEGLGNQKKLREESEAAIFEM
SR12b           TSLENDLPKLYDMVKTEGQDREDNDNGTLRRAGDEIKRLNEGLGNQKKLREESEAAIFEM
SR12c           TSLENDLPKLYDMVKTEGQDREDSDNGTLRRAGDEIKRLNEGLGNQKKLREESEAAIFEM
SRG12d          TSLENDLPKLYDMVKTEGQDREDSDNGTLRRAGDEIKRLNEGLGNQKKLREESEAAIFEM
SR13d           TSLENDLPKLYDMVKSEGQDREDSDNGTLRRSGDEIRRLNEGLTNQKNLREESETAIFEM
SR13c           TSLENDLPKLYDMVKTEGQDREDNDNGTLRRSGEEIRRLNEGLTNQKKLREESETAIFEM
SR7a            QCLETDLPRLHEAIKTEVAEREEMDSNISRKMNEELSKLNQLLVQEKVNRQESEQAIFDM
SR7b            QCLETDLPRLHEAIKTEVAEREEMDSNISRKMNEELSKLNQLLVQEKVNRQESEQAIFDM
SR9             QGLQNDLPKIQEAIRNESNERDESDQNVMKSITDELVKLSNLINVEKRNRDESEQSIFEM
SR8a            QGLQNDLPKIQEAIRNEANERDESDQNVMKSITDELVKLSNLINVEKRNRDESEQSIFEM
SR8b            QGLQNDLPKIQEAIRNEANERDESDQNVMKSITDELVKLSNLINVEKRNRDESEQSIFEM
SR10a           QGLQNDLPKIQEAIREEANERDESDQNVMKSITDELVKLSNLINVEKRNRDESEQSIFEM
SR10b           QGLQNDLPKIQEAIREEANERDESDQNVMKSITDELVKLSSLINVEKRNRDESEQSIFEM
SR10c           QGLQNDLPKIQEAIREEANERDEADQNVMKSITDELVKLSNLINVEKRNRDESEQSIFEM
SR10d           QGLQNDLPKIQEAIREEANERDEADQNVMKSITDELVKLSNLINVEKRNRDESEQSIFEM
                  :   : .: : :  *  :* : :        :.:  *   .  :.  *:  :  :.
```

```
SR5a            LREVQFRITELLKRNRKERETTEQEMVSL----------VETVIE---------------
SR5b            LREVQFRITELLKRNRKERETTEQEMVSL----------VETVIE---------------
SR4             LREISGRITEALKKTRTEREKTEETLVQL----------VEKVVE---------------
SR6d            LREISGRIQEQLRKTRGEREKTEETLVQL----------VEKVIE---------------
SR6c            LREISGRIQEQLRKTRGEREKTEETLVQL----------VEKVIE---------------
SR6a            LREISGRIQEQLKKTRTEREKTEETLVQL----------VEKVIE---------------
SR6d            LREISGRIQEQLKKTRNEREKTEETLVQL----------VEKVIE---------------
SR3             LREISIRIQEQLRKNKGEREKTEETLVTL----------VEKVIE---------------
SR2             LREISNRIQEQLRKTRGEREKTEETLVTL----------VEKVIE---------------
SR1a            LDEMNHYLNGELQSEKNEREATEESLINL----------IDQTCN---------------
SR1b            LDEMNHYLNSELTSEKNEREATEESLINL----------IDQTCN---------------
SR13a           LKDLVSRVKSEIEEEKKLREESQESLLGL----------LEDAAN---------------
SR13b           LKDLVSRVKSEIEEEKKLREESQESLLGL----------LEDAAN---------------
SR11a           LKDLVSRVKSEIEEEKKLREESQENLLGL----------LEDACN---------------
SR11b           LKDLVSRVKSEIEEEKKLREESQENLLGL----------LEDACN---------------
SR12a           LKDLVSRVKSEIEEEKKLREESQESLLGL----------LEDACN---------------
SR12b           LKDLVSRVKSEIEEEKKLREESQESLLGL----------LEDACN---------------
SR12c           LKDLVSRVKSEIEEEKKLREESQESLLGL----------LEDACN---------------
SRG12d          LKDLVSRVKSEIEEEKKLREESQESLLGL----------LEDACN---------------
SR13d           LKDLVSRVKSEIEEERKLREESQESLLGL----------LEDAAN---------------
SR13c           LKDLVSRVKSEIEEEKKLREESQESLLGL----------LEDAAN---------------
SR7a            LKDVVNRIKTEVDNEKKQRESTEETLLAL----------LEDTCN---------------
SR7b            LKDVVNRIKTEVDNEKKQRESTEETLLAL----------LEDTCN---------------
SR9             LKDIVNRVKVELDQEKKTREQSEEHLLSL----------LEDTCN---------------
SR8a            LKDIVNRVKVELDQEKKTREQSEEHLLSL----------LEDTCN---------------
SR8b            LKDIVNRVKVELDQEKKTREQSEEHLLSL----------LEDTCN---------------
SR10a           LKDIVNRVKVELDQEKRTREQSEEHLLSL----------LEDTCN---------------
SR10b           LKDIVNRVKVELDQEKRTREQSEEHLLSL----------LEDTCN---------------
SR10c           LKDIVNRVKVELDQEKRTREQSEEHLLSL----------LEDTCN---------------
SR10d           LKDIVNRVKVELDQEKRTREQSEEHLLSLCWKTHVINLVLQQIYDIYIYIMNYSLVKVSI
                * ::   :   :   .  ** ::: :: *          ::   :
```

```
SR5a            ----KIKIEMLEMNM
SR5b            ----KIKIEMLEMNM
SR4             ----KIKREMLEMNL
SR6d            ----KLKREMLEMNL
SR6c            ----KLKREMLEMNL
SR6a            ----KLKREMLEMNL
SR6d            ----KLKREMLEMNL
SR3             ----KLKREMLEMNL
SR2             ----KLKREMLEMNL
SR1a            ----RVENSLRK---
SR1b            ----RVENSLRK---
SR13a           ----KIYRAAKD---
SR13b           ----KIYKAAKD---
SR11a           ----KIYRAAKD---
SR11b           ----KIYRAAKD---
SR12a           ----KIYRAAKD---
SR12b           ----KIYRAAKD---
SR12c           ----KIYRAAKD---
SRG12d          ----KIYRAAKD---
SR13d           ----KIYRAAKD---
SR13c           ----KIYRAAKD---
SR7a            ----KVNAAQLA---
SR7b            ----KVNAAQLA---
SR9             ----KLSIAANL---
SR8a            ----KLSIAANL---
SR8b            ----KLSIAANL---
SR10a           ----KLSIAANL---
SR10b           ----KLSIAANL---
SR10c           ----KLSIAANL---
SR10d           FITLGIQHTVLNYTT
```

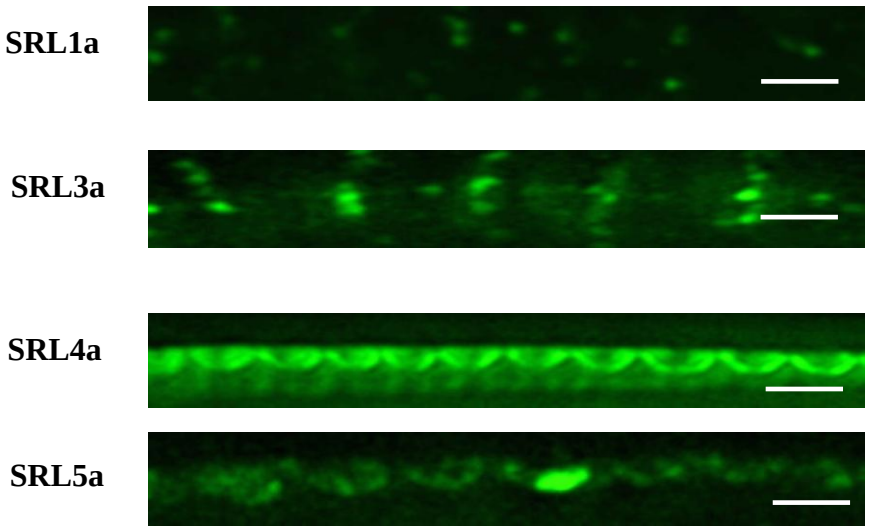
2.S3: The phylogenetic relationships among the *SR* and *SRL* genes in *Paramecium*.

Evolutionary analysis of the 30 *SR* genes and 11 *SRL* genes were conducted in MEGA6 (Tamura et al., 2013). The evolutionary history was inferred using the Neighbor-Joining method (Saitun and Nei, 1987). The analysis involved all 41 nucleotide sequences. The final dataset includes 475 positions and eliminates all positions having the gap and missing data sets.


SR 10d
SR 10c
SR 10b
SR 10a
SR 9
SR 8b
SR 8a
SR 7b
SR 7a
SR 1a
SR 1b
SR 5a
SR 5b
SR 4
SR 2
SR 3
SR 6a
SR 6b
SR 6c
SR 6d
SR 13d
SR 13c
SR 13b
SR 13a
SR 12d
SR 12c
SR 12b
SR 12a
SR 11b
SR 11a
SRL 1a
SRL 1b
SRL 2
SRL 3a
SRL 3b
SRL 3c
SRL 4a
SRL 4b
SRL 5a
SRL 5b
SRL 5c

2.S4: SRL proteins without the conserved domains of the *Chlamydomonas* SF-assemblin protein have different locations than SRs.

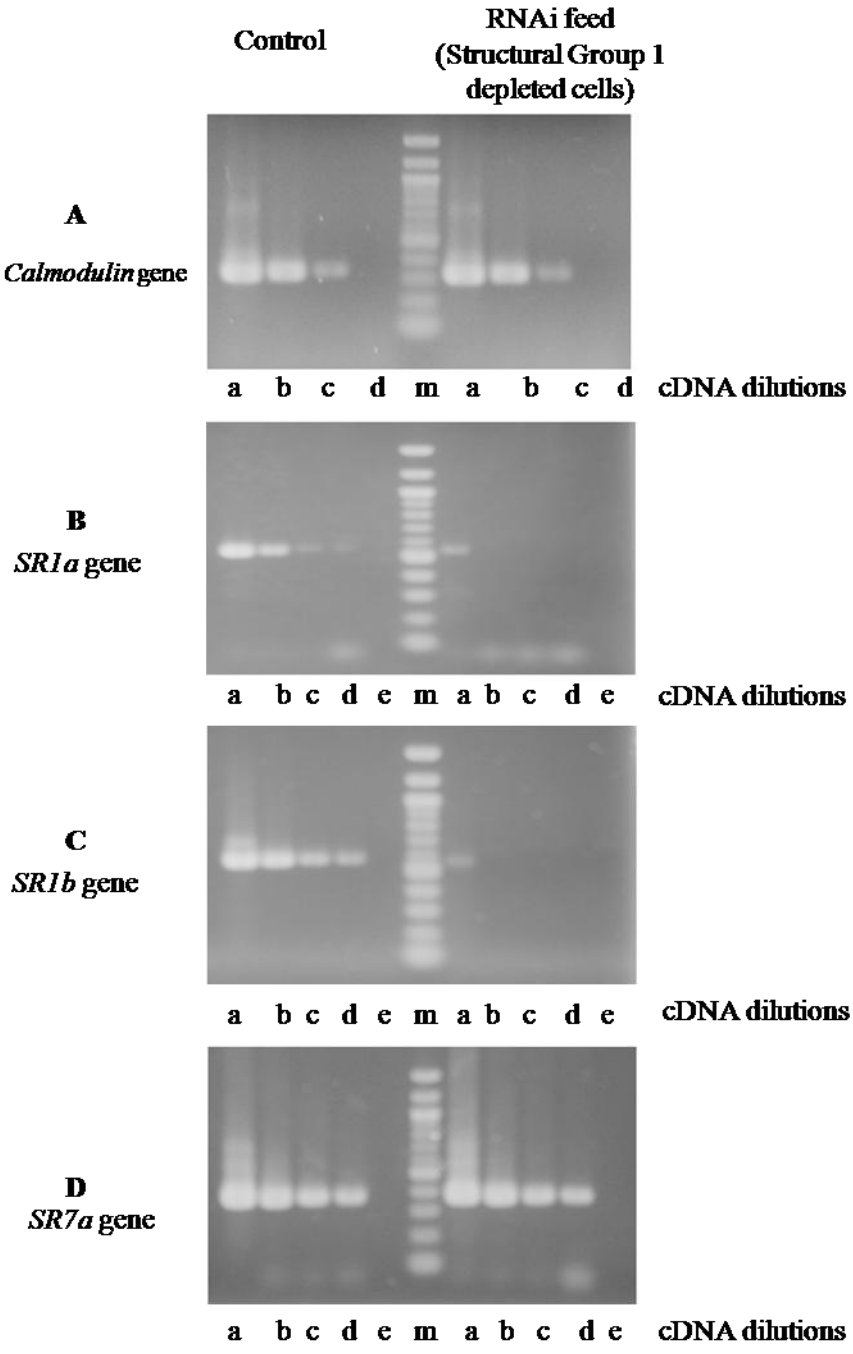
The Panel shows the location of FLAG-SRL protein expressing cells. Both FLAG-SRL4a and FLAG-SRL5a are located in the epiplasm of the *Paramecium* cell. The FLAG-SRL1a is located intracellularly where as FLAG-SRL3a is located in the basal bodies and throughout the cilia.


SRL1a
SRL3a
SRL4a
SRL5a

2.S5: The RT-PCR analysis showing the depletion of mRNA of targeted SR Structural Group.

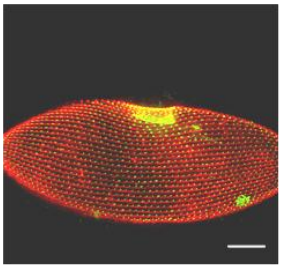
Panel A shows a representative image of an agarose gel demonstrating the level of m-RNA for *calmodulin* gene using gene specific primers in control L4440 fed cells (right hand side) or Structural Group 1 (Paralog Group: SR 1) depleted cells (left hand side). In Panel A, m indicates the 100 bp DNA Ladder and a, b and c indicates 1:10, 1:100 and 1:1000 dilution of c-DNA which was made using reverse transcriptase enzyme after extraction of total m-RNA from the control cells or the RNAi (SR Structural Group 1) cells. The d indicates 1:10 dilution of cDNA which was made without using reverse transcriptase enzyme after extraction of total mRNA from the control cells or the Structural Group 1 depleted cells.

Panel B, C, and D show representative image of an agarose gel demonstrating the level of mRNA for *SR1a* gene (targeted gene), *SR1b* gene (targeted gene) and *SR7a* gene (nontargeted gene) respectively using gene specific primers in control L4440 fed cells (right hand side) or Structural Group 1 (Paralog Group: SR 1) depleted cells (left hand side). In panels, m indicate the 100 bp DNA Ladder and a, b, c and d indicate 1:10, 1:100, 1:200 and 1:1000 dilution of c-DNA which was made using reverse transcriptase enzyme after extraction of total mRNA from the control cells or the test cells. In panels, e indicate 1:10 dilution of cDNA which was made without using reverse transcriptase enzyme after extraction of total mRNA from the control cells or the RNAi (SR Structural Group 1) depleted cells.

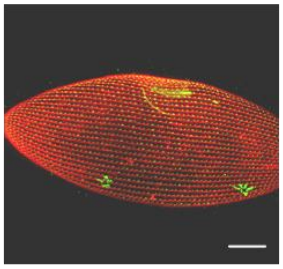

Control
RNAi feed
(Structural Group 1
depleted cells)
A
Calmodulin gene
a b c d m a b c d cDNA dilutions
B
SR1a gene
a b c d e m a b c d e cDNA dilutions
C
SR1b gene
a b c d e m a b c d e cDNA dilutions
D
SR7a gene
a b c d e m a b c d e cDNA dilutions

2.S6: The whole cell image demonstrating the phenotype of the basal body units and associated SRs in SR Paralog Group or SR Structural Group depleted cells.

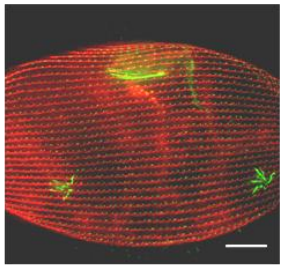
Panels A-K show the images of whole cell showing the basal body rows (green) and SRs (red) in the SR Paralog Group SR 2, SR 3, SR 4, SR 9, SR 5, SR 6, SR 8, SR 10, SR 11, SR 12 or SR 13 depleted cells. Panels L-O show the images of whole cell showing the basal body rows (green) and SRs (red) in the SR Structural Group 1 (Paralog Group: SR 1), Structural Group 5 (Paralog Group: SR 2-SR 6), Structural Group 4 (Paralog Group: SR 8-SR 10) or Structural Group 3 (Paralog Group: SR 11-SR 13) depleted cells. Depletion of Paralog Group (Panels A-K) does not affect basal body row alignment and the highly ordered organization of the SRs. Depletion of Structural Groups (Panels L-O) results in misaligned basal body rows with the disordered organization of the SRs. Scale bars are 15μm. Panels G, H and I have been rotated for easier comparison.



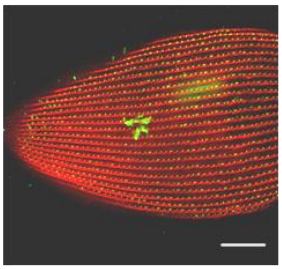
A (SR 2)



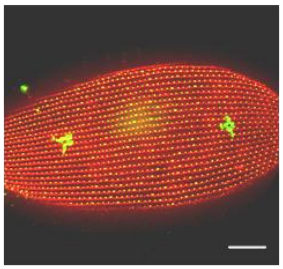
B (SR 3)



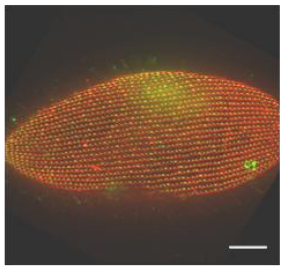
C (SR 4)



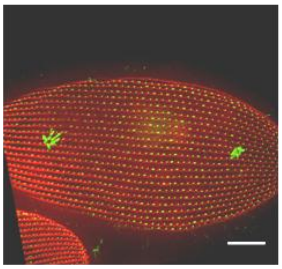
D (SR 9)



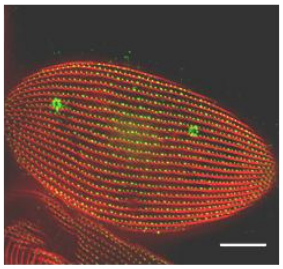
E (SR 5)



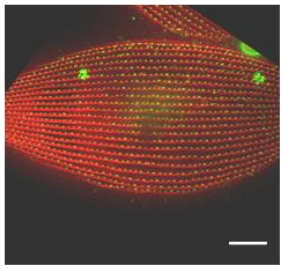
F (SR 6)



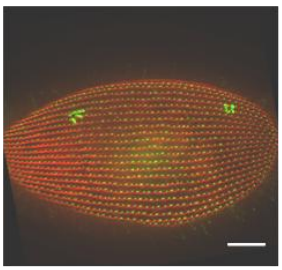
G (SR 8)



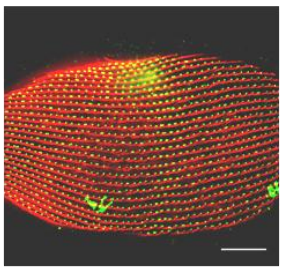
H (SR 10)



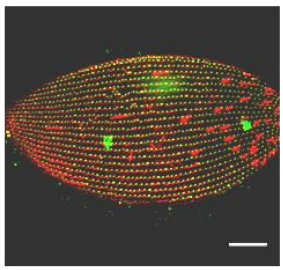
I (SR 11)



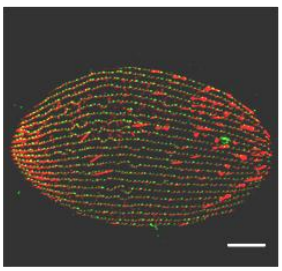
J (SR 12)



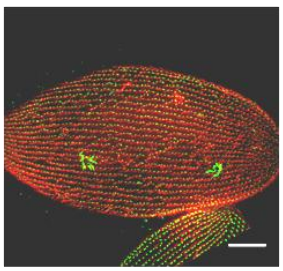
K (SR 13)



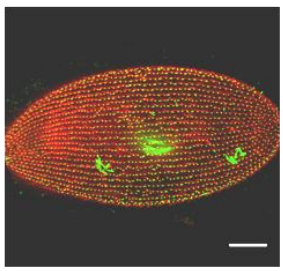
L (SR 1)



M (SR 2-SR 6)



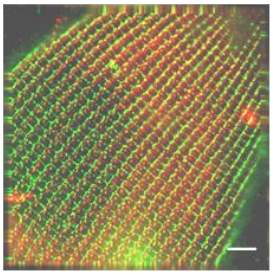
N (SR 8-SR 10)



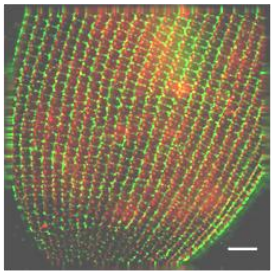
O (SR 11-SR 13)

2.S7: The whole cell images demonstrating the shape of cortical unit and the SRs in SR Paralog Group or SR Structural Group depleted cells.

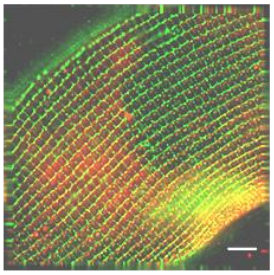
Panels A-K show the images of whole cell showing the cortical unit (green) and the SRs (red) in the SR Paralog Group SR 2, SR 3, SR 4, SR 9, SR 5, SR 6, SR 8, SR 10, SR 11, SR 12 or SR 13 depleted cells. Panels L-O show the images of whole cell showing the cortical unit (green) and SRs (red) in the SR Structural Group 1 (Paralog Group: SR 1), Structural Group 5 (Paralog Group: SR 2-SR 6), Structural Group 4 (Paralog Group: SR 8-SR 10) or Structural Group 3 (Paralog Group: SR 11-SR 13) depleted cells. Depletion of Paralog Group (Panels A-K) does not affect the highly organized cortical unit rows and the highly ordered organization of SRs. Depletion of Structural Groups (Panels L-O) results in distorted cortical units with abnormal SRs. Scale bars are 10 μm.



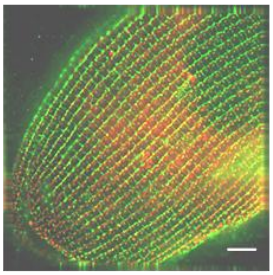
A (SR 2)



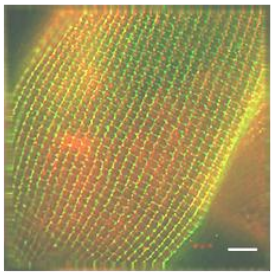
B (SR 3)



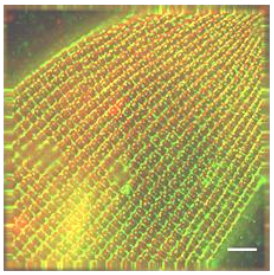
C (SR 4)



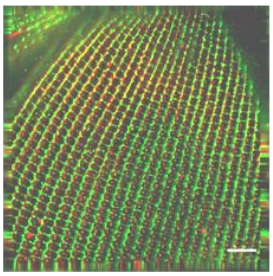
D (SR 9)



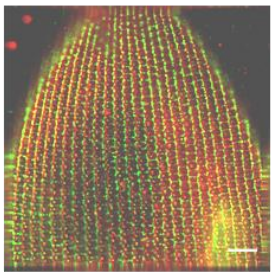
E (SR 5)



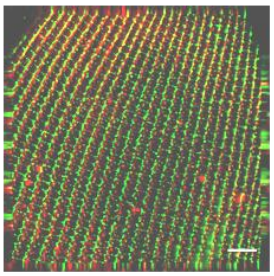
F (SR 6)



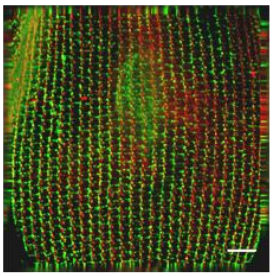
G (SR 8)



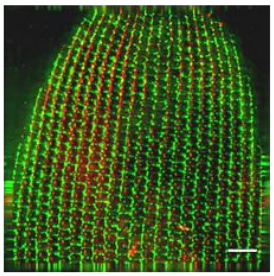
H (SR 10)



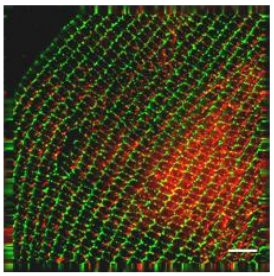
I (SR 11)



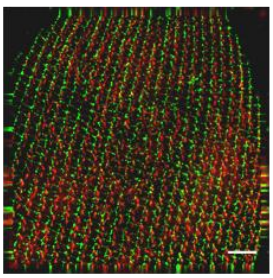
J (SR 12)



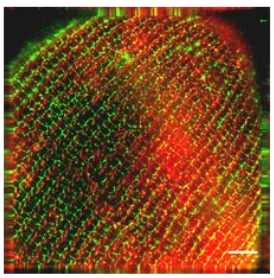
K (SR 12)



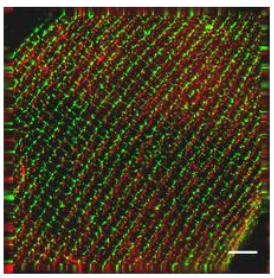
L (SR 1)



M (SR 2-SR 6)



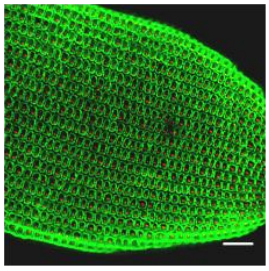
N (SR 8-SR 10)



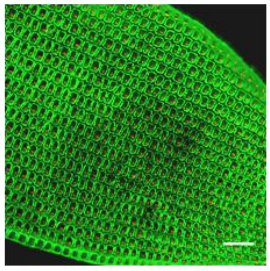
O (SR 11-SR 13)

2.S8: The whole cell images demonstrating the territory of epiplasm unit and the basal body units in SR Paralog Group or SR Structural Group depleted cells.

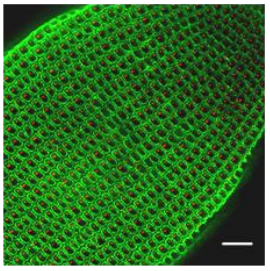
Panels A-K show the images of whole cell showing the epiplasm unit (green) and the basal body units (red) in the SR Paralog Group SR 2, SR 3, SR 4, SR 9, SR 5, SR 6, SR 8, SR 10, SR 11, SR 12 or SR 13 depleted cells. Panels L-O show the images of whole cell showing the epiplasm unit (green) and SRs (red) in the SR Structural Group 1 (Paralog Group: SR 1), Structural Group 5 (Paralog Group: SR 2-SR 6), Structural Group 4 (Paralog Group: SR 8-SR 10) or Structural Group 3 (Paralog Group: SR 11-SR 13) depleted cells. Depletion of Paralog Group (Panels A-K) does not affect the highly organized epiplasm unit and the straight basal body rows on the cell surface. Depletion of Structural Groups (Panels L-O) results in irregular shaped epiplasm unit with misaligned basal body rows. Scale bars are 10 μm.



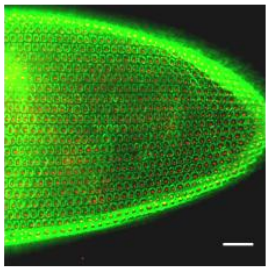
A (SR 2)



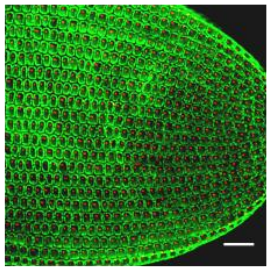
B (SR 3)



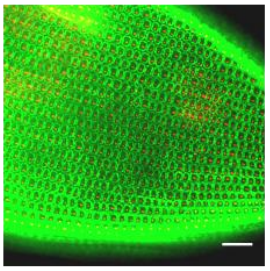
C (SR 4)



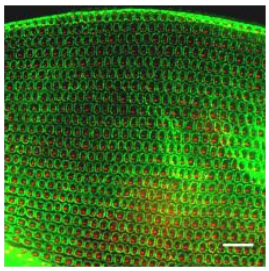
D (SR 9)



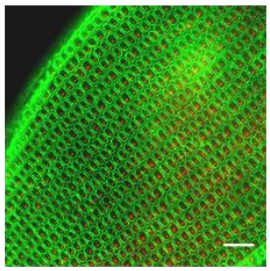
E (SR 5)



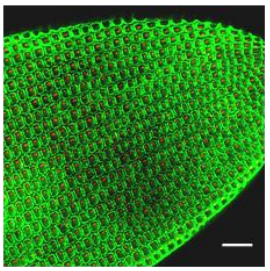
F (SR 6)



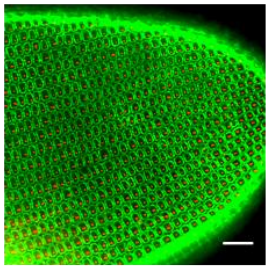
G (SR 8)



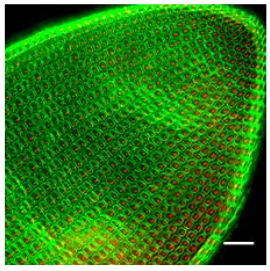
H (SR 10)



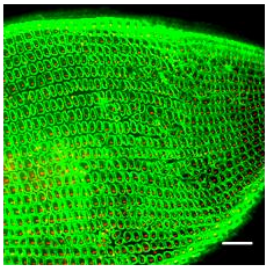
I (SR 11)



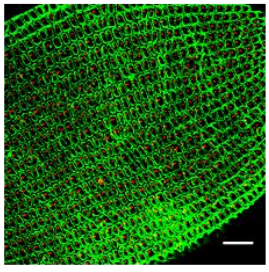
J (SR 12)



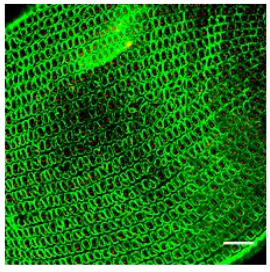
K (SR 13)



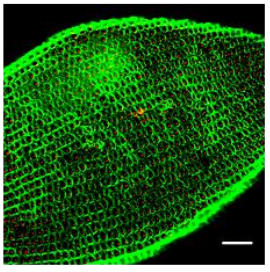
L (SR 1)



M (SR 2-SR 6)



N (SR 8-SR 10)

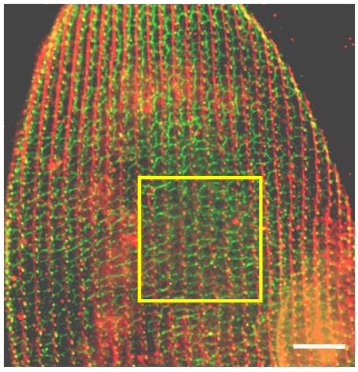


O (SR 11-SR 13)

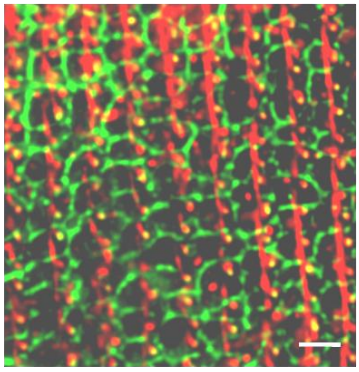
2.S9: Depletion of SR Structural Group affects the association between ICL complex network and SRs along with misaligned basal body rows.

Panels A and C show the ICL network in the SR Paralog Group SR 10 and SR Structural Group 1 (SR Paralog Group: SR 1) depleted cell respectively. Both the cells are injected to express FLAG-Ptcen15. Anti-Flag staining of the cells decorate ICL network (green). The basal body units (red) and associated SRs (red) are stained with anti-centrin and anti-SR respectively. The yellow box in each image is enlarged (B and D) to demonstrate ICL network along with basal body units (red) and SRs (red). All images are of the dorsal surface of the cell. Panel B shows close association of ICL network and basal body units along with SRs in SR Paralog Group SR 10 depleted cell. Panel D shows the association is somehow missing in between the ICL network and basal body units with abnormal SRs in SR Structural Group 1 depleted cell. Scale bars are 10μm (A and C) and 3μm (B and D)**.**

SR 10 depleted cells

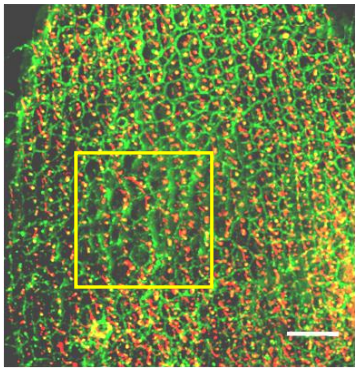


A

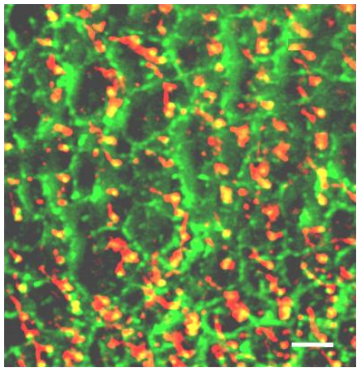


B

SR Structural Group 1 depleted cells



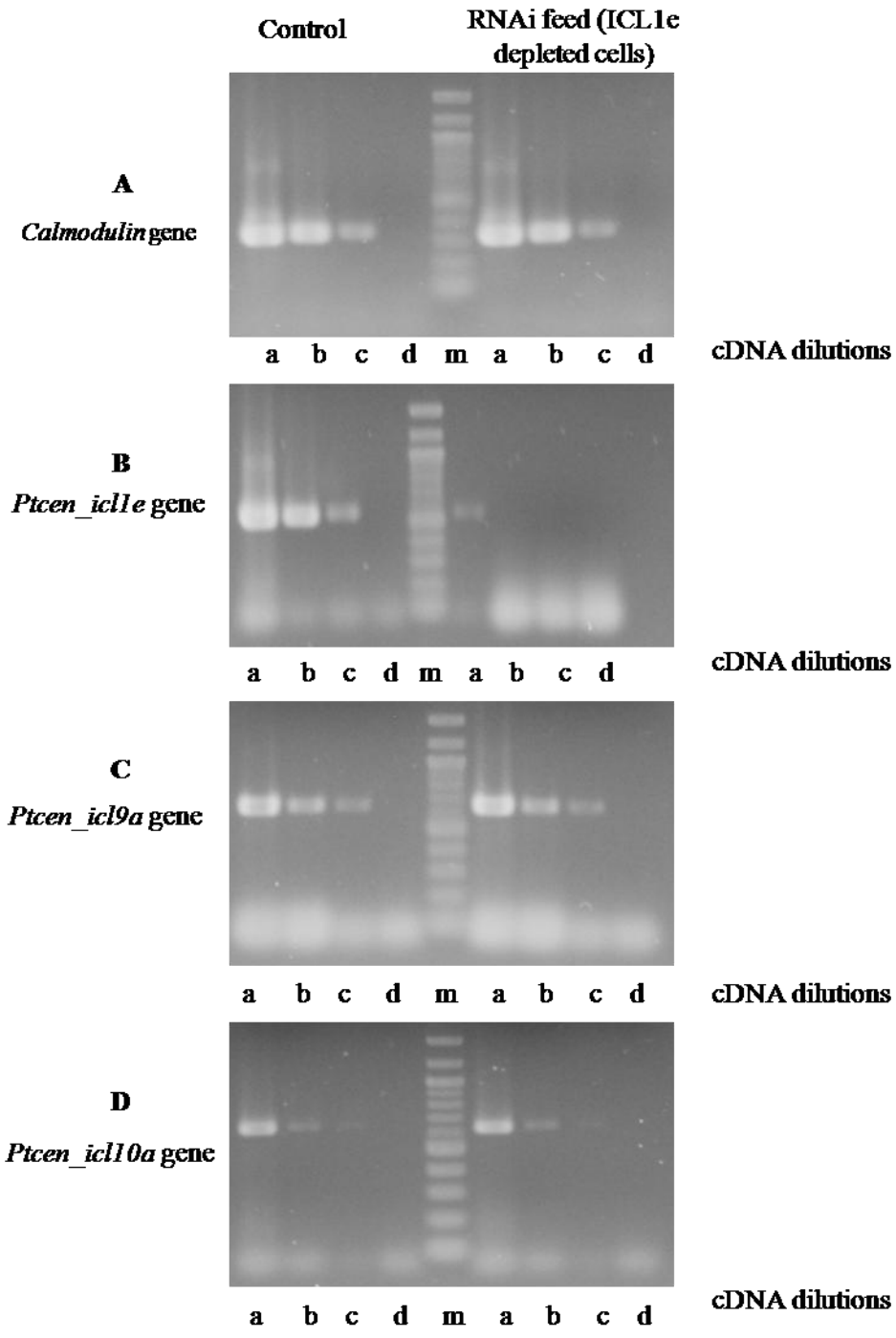
C



D

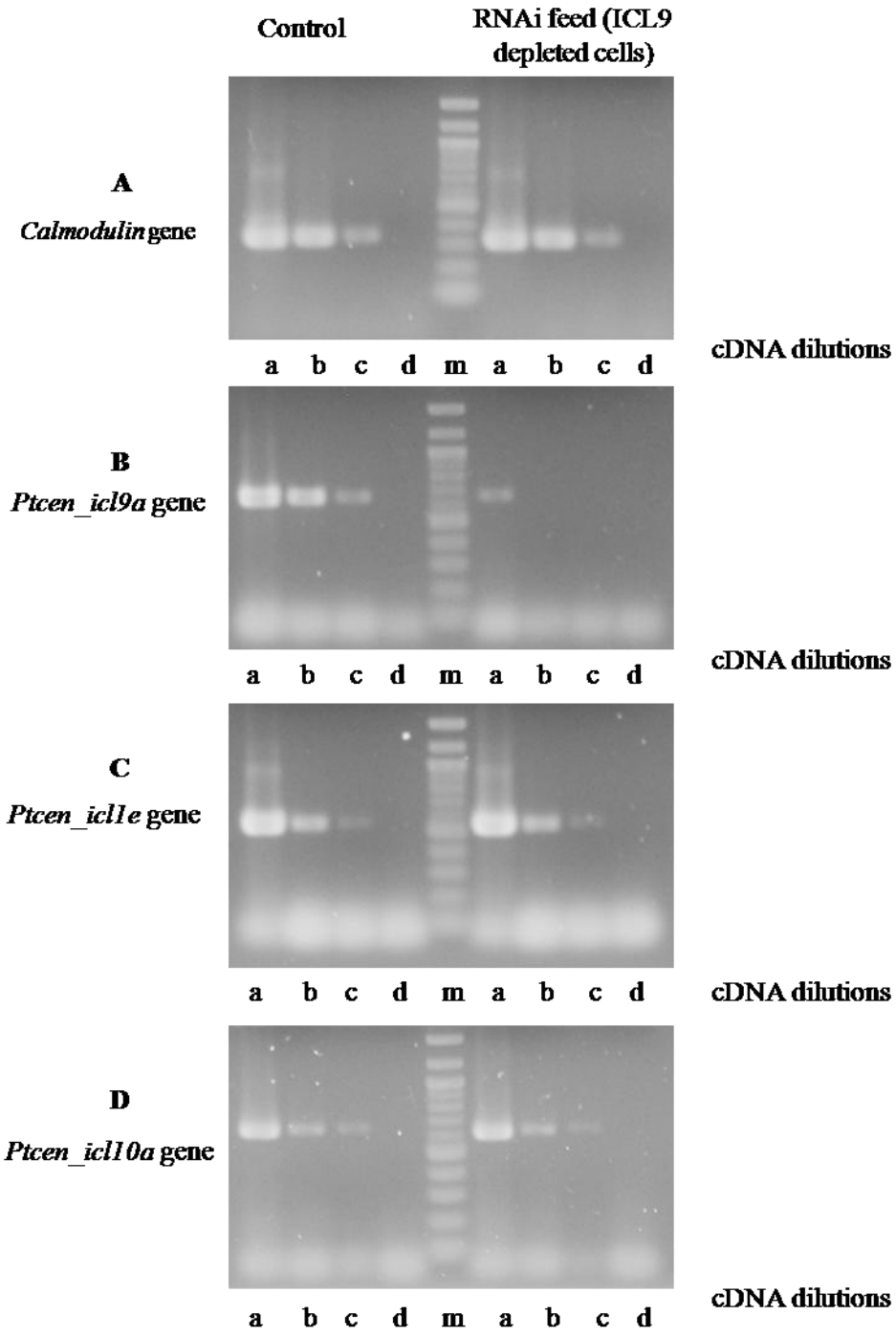
2.S10: The RT-PCR analysis showing the depletion of mRNA of targeted ICL1e subfamily.

Panel A, B, C and D show a representative image of an agarose gel demonstrating the level of mRNA for *calmodulin* gene, *Ptcen_icl1e* gene (targeted gene), *Ptcen_icl9a* gene (nontargeted gene) and *Ptcen_icl10a* gene (nontargeted gene) using gene specific primers in control L4440 fed cells (right hand side) or ICL1e subfamily depleted cells (left hand side). In all panels , m indicate the 100 bp DNA Ladder and a, b and c indicate 1:10, 1:100 and 1:1000 dilution of cDNA which was made using reverse transcriptase enzyme after extraction of total mRNA from the control cells or the ICL1e subfamily depleted cells. The d indicates 1:10 dilution of cDNA which was made without using reverse transcriptase enzyme after extraction of total mRNA from the control cells or the ICL1e subfamily depleted cells.


Control
RNAi feed (ICL1e depleted cells)
A
Calmodulin gene
a b c d m a b c d
cDNA dilutions
B
Ptcen_icl1e gene
a b c d m a b c d
cDNA dilutions
C
Ptcen_icl9a gene
a b c d m a b c d
cDNA dilutions
D
Ptcen_icl10a gene
a b c d m a b c d
cDNA dilutions

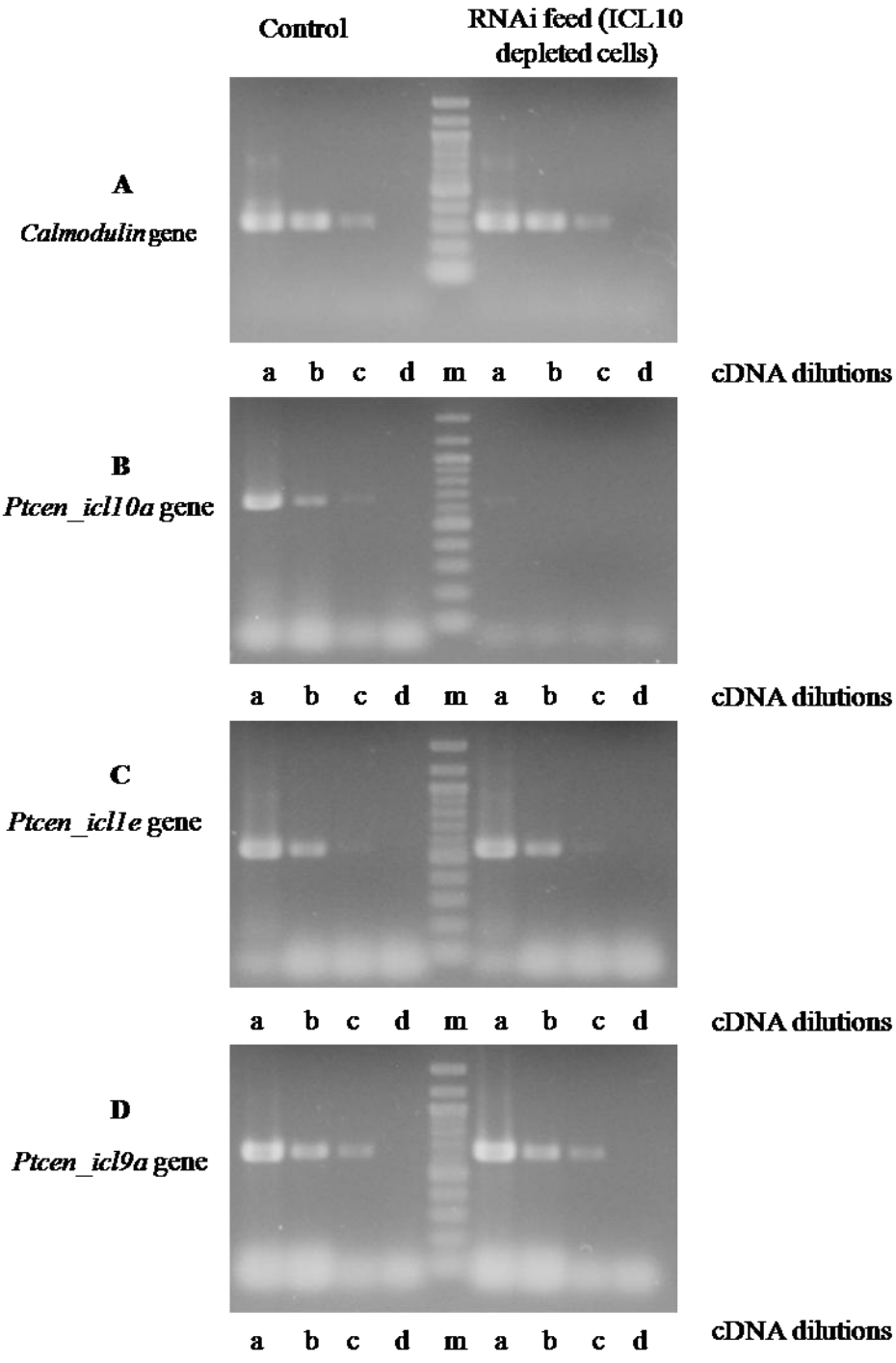
2.S11: The RT-PCR analysis showing the depletion of mRNA of targeted ICL9 subfamily.

Panels A, B, C and D shows the representative image of an agarose gel demonstrating the level of mRNA for *calmodulin* gene, *PtCen_icl9a* gene (targeted gene), *Ptcen_icl1e* gene (nontargeted gene) and *Ptcen_icl10a* gene (nontargeted gene) using gene specific primers in control L4440 fed cells (right hand side) or ICL9 subfamily depleted cells (left hand side). In Panels, m indicates the 100 bp DNA Ladder and a, b and c indicates 1:10, 1:100 and 1:1000 dilution of cDNA which was made using reverse transcriptase enzyme after extraction of total mRNA from the control cells or the test cells. The d indicates 1:10 dilution of cDNA which was made without using reverse transcriptase enzyme after extraction of total mRNA from the control cells or the ICL9 subfamily depleted cells.


Control
RNAi feed (ICL9 depleted cells)
A
Calmodulin gene
a b c d m a b c d
cDNA dilutions
B
Ptcen_icl9a gene
a b c d m a b c d
cDNA dilutions
C
Ptcen_icl1e gene
a b c d m a b c d
cDNA dilutions
D
Ptcen_icl10a gene
a b c d m a b c d
cDNA dilutions

2.S12: The RT-PCR analysis showing the depletion of mRNA of targeted ICL10 subfamily.

Panels A, B, C and D show the representative image of an agarose gel demonstrating the level of mRNA for *calmodulin* gene, *PtCen_icl10a* gene (targeted gene), *Ptcen_icl1e* gene (nontargeted gene) and *Ptcen_icl9a* gene (nontargeted gene) using gene specific primers in control L4440 fed cells (right hand side) or ICL10 subfamily depleted cells (left hand side). In panels m indicates the 100 bp DNA Ladder, a, b and c indicates 1:10, 1:100 and 1:1000 dilution of cDNA which was made using reverse transcriptase enzyme after extraction of total m-RNA from the control cells or the test cells. The d indicates 1:10 dilution of cDNA which was made without using reverse transcriptase enzyme after extraction of total mRNA from the control cells or the ICL10 subfamily depleted cells.


Control
RNAi feed (ICL10 depleted cells)
A
Calmodulin gene
a b c d m a b c d
cDNA dilutions
B
Ptcen_icl10a gene
a b c d m a b c d
cDNA dilutions
C
Ptcen_icl1e gene
a b c d m a b c d
cDNA dilutions
D
Ptcen_icl9a gene
a b c d m a b c d
cDNA dilutions

2.S13: Proteins found by LC-MS/MS in the 30% Optiprep fraction in silver stained gels that exclude the area of the gel shown in red box in the Fig.2.6B.

Protein group	Protein Name	Total peptide sequences identified	Unique peptide sequence
PtCenBP1/ CenBP1	-	21	-
Beta_tubulin	tub_betaPT1	16	0
	tub_betaPT1		0
ICL 1e	ICL1e	6	0
	ICL1g		0
	Ptcen8		0
	Ptcen15		0
	Ptcen10		0
	Ptcen12		1
	Ptcen18		0
Ribosomal protein	GSPATP00004608001	5	-
Alpha_tubulin	tub_alphaPT3	4	0
	tub_alphaPT4		0
	tub_alphaPT2		0
	tub_alphaPT1		0
	tub_alphaPT1		0
	tub_alphaPT8		0
Ribosomal protein	GSPATG00019021001	6	0
	GSPATG00004327001		0
Ribosomal protein	GSPATG00004608001	5	-
Heat shock protein	GSPATG00023365001	4	0
	GSPATG00021676001		0
	GSPATG00030693001		0
	GSPATG00026329001		0
Heat shock protein	GSPATG00026329001	5	0
	GSPATG00030693001		0
	GSPATG00021676001		0
	GSPATG00023365001		0
Ribosomal protein	GSPATG00033610001	9	3
	GSPATG00023827001		0
Ribosomal protein	GSPATG00012801001	7	1
	GSPATG00013209001		0
Heat shock protein	GSPATG00018573001	9	2
	GSPATG00017143001		0

2.S14: SR proteins and other proteins found by LC-MS/MS in SR preparation isolated from SR Structural Group 1 (SR 1) depleted cell (data combined from 3 experiments).

	SR protein	Total peptide sequences identified	Unique Peptide
SR 2-6 (Structural Group 5)	SR 2	30	2
	SR 3		0
	SR 4		2
	SR 5a		1
	SR 5b		1
	SR 6a		0
	SR 6b		0
	SR 6c		1
	SR 6d		0
SR 7 (Structural Group 2)	SR 7a	7	0
	SR 7b		0
SR 8-10 (Structural Group 3)	SR 8a	15	1
	SR 8b		0
	SR 9		1
	SR 10a		0
	SR 10b		1
	SR 10c		0
	SR 10d		0
SR 11-13 (Structural Group 4)	SR 11a	27	1
	SR 11b		1
	SR 12a		0
	SR 12b		0
	SR 12c		0
	SR 12d		0
	SR 13a		2
	SR 13b		1
	SR 13c		2
	SR 13d		0
Beta_tubulin	tub_betaPT1	11	0
	tub_betaPT1		0
Ribosomal protein	GSPATP00004608001	3	-
Alpha_tubulin	tub_alphaPT3	6	0
	tub_alphaPT4		0
	tub_alphaPT2		0
	tub_alphaPT1		0

	tub_alphaPT1		0
	tub_alphaPT8		0
Ribosomal protein	GSPATG00019021001	2	0
	GSPATG00004327001		0
Heat shock protein	GSPATG00023365001	2	0
	GSPATG00021676001		0
	GSPATG00030693001		0
	GSPATG00026329001		0
Heat shock protein	GSPATG00018573001	3	2
	GSPATG00017143001		0
Ribosomal protein	GSPATG00033610001	4	3
	GSPATG00023827001		0
Heat shock protein	GSPATG00018573001	3	2
	GSPATG00017143001		0

2.S15: SR proteins and other proteins found by LC-MS/MS in SR preparation isolated from SR Structural Group 2 (SR 7) depleted cell. (data combined from 3 experiments).

	SR protein	Total peptide sequences identified	Unique Peptide
SR 1 (Structural Group 1)	SR 1a	10	3
	SR 1b		5
SR 2-6 (Structural Group 5)	SR 2	25	2
	SR 3		0
	SR 4		2
	SR 5a		1
	SR 5b		1
	SR 6a		0
	SR 6b		0
	SR 6c		1
	SR 6d		0
SR 8-10 (Structural Group 3)	SR 8a	12	1
	SR 8b		0
	SR 9		1
	SR 10a		0
	SR 10b		1
	SR 10c		0
	SR 10d		0
SR 11-13 (Structural Group 4)	SR 11a	13	1
	SR 11b		1
	SR 12a		0
	SR 12b		0
	SR 12c		0
	SR 12d		0
	SR 13a		2
	SR 13b		1
	SR 13c		2
	SR 13d		0
Beta_tubulin	tub_betaPT1	11	0
	tub_betaPT1		0
Ribosomal protein	GSPATP00004608001	3	-
Alpha_tubulin	tub_alphaPT3	6	0
	tub_alphaPT4		0
	tub_alphaPT2		0
	tub_alphaPT1		0

	tub_alphaPT1		0
	tub_alphaPT8		0
Ribosomal protein	GSPATG00019021001	2	0
	GSPATG00004327001		0
Heat shock protein	GSPATG00023365001	6	0
	GSPATG00021676001		0
	GSPATG00030693001		0
	GSPATG00026329001		0
Ribosomal protein	GSPATG00033610001	4	3
	GSPATG00023827001		0
Heat shock protein	GSPATG00018573001	3	2
	GSPATG00017143001		0
Ribosomal protein	GSPATG00033610001	6	3
	GSPATG00023827001		0
Heat shock protein	GSPATG00026329001	5	0
	GSPATG00030693001		0
	GSPATG00021676001		0
	GSPATG00023365001		0

Chapter 3: Depletion of MKS5 causes loss of cilia but does not affect the basal body rows straight alignment

Introduction:

Cilia are slender, microscopic, whip-like conserved structures found on the cell surface. In mammals, almost all cells have a primary cilium or are multi-ciliated. This primordial organelle performs an important role in the process of signal transduction, cell movement and development. Since cilia perform very important functions, abnormal function or structure of the cilia can contribute to a group of autosomal or x-linked diseases, collectively termed as ciliopathies (Fliegauf et al., 2007; Afzelius, 2004; Nakamura et al., 2006). Ciliopathies can be characterized by multi-cystic kidneys, central nervous system (CNS) malformations, polydactyly, respiratory illness, liver malformations, retinal degeneration, diabetes, obesity and skeletal dysplasias (Afzelius, 2004; Badano et al.,2005; Badano et al., 2006).

A cilium has a unique structure and organization as a cell organelle. While the ciliary membrane is a lipid bi-layer membrane that is contiguous with the cell body membrane, the lipid and protein composition of these two membranes differ remarkably (Rohatgi and Snell, 2010). The basal body at the base of the cilium plays important role to anchor the cilium on the cell surface. It has nine-fold symmetry with microtubule triplets that form a characteristic cartwheel structure. The basal body structure also organizes the microtubules of the ciliary axoneme to give characteristic ′9+2′ or ′9+0′ structure. The distal end of the basal body where microtubule triplets convert to microtubule doublets to organize the core of ciliary axoneme is the area of the transition zone (TZ). TZ acts as a molecular sieve that controls the entry and exit of different

proteins in the cilia (Czarnecki and Shah, 2012; Kobayashi and Dynlacht, 2011). The ciliary compartment does not have ribosome or the protein synthesis machinery, making the biogenesis, function and maintenance of cilia dependent upon an exclusive bi-directional transport process known as intraflagellar transport (IFT).

Many of the reported ciliopathy genes are associated with TZ and basal body of cilia (Waters and Beales, 2011). The TZ of ciliary basal body includes several multi-protein complexes such as NPHP 1-4-8, MKS/B9, NPHP 5-6 and inversion compartment complex. The MKS multi-protein complex consists of MKS1, MKS2 (TMEM216), MKS3 (Meckelin/TMEM67), MKS6 (CC2D2A), MKS8 (TCTN2), TCTN-1, MKS9 (B9D1) and MKS10 (B9D2), The NPHP 1-4-8 multi-protein complex consists of MKS5 (RPGRIP1), NPHP1 and NPHP4 proteins. The MKS5 protein of the NPHP 1-4-8 multi-protein complex plays an important role in the formation of the TZ of the cilia. It interacts with other proteins and influences the formation of MKS/B9 multiprotein complex (Czarnecki and Shah, 2012).

Phenotypic features of the ciliopathy diseases can be heterogeneous. Some closely related ciliopathy diseases such as Joubert syndrome and related disorders (JSRD), Meckel Gruber syndrome (MKS) and Nephronophthisis share the common clinical features that include renal dysplasia, occipital encephalocele and polydactyly (Salonen and Norio1984; Salonen and Paavola, 1998). The proteins such as MKS3, CEP290, MKS5, MKS6, and MKS2 have reported being mutated in these ciliopathy diseases (Smith et al., 2006; Kyttala et al., 2006; Tallila et al., 2008; Valente et al., 2006). Therefore, to understand the heterogeneous nature of the ciliopathy diseases, it is

important to investigate the diverse functions of the ciliopathy proteins in the tissue or cell-specific manner.

The ciliate *Paramecium* is an excellent model for studying the cilia especially the biology of the basal body, centrioles and ciliopathy genes. The number of characteristics that make *Paramecium* an attractive model include high number of basal bodies (~4000 basal bodies per cell), large number of cilia, the sequenced and annotated genome, the presence of mammalian orthologs of ciliopathy genes, and, finally, the ease of applying various genetic and biochemical approaches. (Vincensini et al., 2011).

In *Paramecium* like other ciliates, basal body structure has the core ′9+2′ structure. The cell cortex is divided into highly organized hexagonal shaped cortical units. Either a single or pair of basal bodies can be found in each cortical unit. Each cortical unit contains a fibrous submembranous cytoskeleton structure (epiplasm), which follows the boundary of the cortical unit (Allen, 1971; Damaj et al., 2009; Aubusson-Fleury et al., 2013). In the single basal body cortical unit, the basal body is accompanied by a set of three asymmetrical rootlets. Among the three rootlets, the largest one is the striated rootlet (SR) that projects from the basal body towards the anterior pole of the cell and spans several cortical units. The other two rootlets are the microtubule ribbons called the postciliary rootlet (PCR) that projects towards the posterior pole of the cell and the transverse rootlet (TR) that projects laterally towards the adjacent rows of basal body units. In the paired basal bodies unit, the posterior basal body is accompanied by all the three asymmetrical rootlets and the anterior basal body is only accompanied by the TR rootlet. There is a linker that connects the anterior basal body to the SR rootlet of the posterior basal body (Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003). All the basal

body units (single or paired, along with their associated rootlets) form straight lines of basal bodies from the posterior to the anterior pole of the cell both in dorsal and ventral surfaces, but not in the oral groove area. In the oral groove area, basal body units are much more tightly packed in the gullet structure (Iftode et al., 1989; Iftode and Fleury-Aubusson, 2003).

We previously reported (Picariello et al., 2014) that MKS3 in *Paramecium* is localized in the TZ and functions in maintaining the normal length of cilia and ciliary function (Smith et al., 2006; Dawe et al., 2007). We also observed a previously undefined function of the MKS3 protein in maintaining the organization of basal body alignment in straight rows predominantly at the dorsal midline of the cell (Picariello et al., 2014).The work reported here examines the role of MKS5 protein by RNA interference to determine whether this protein, which is thought to be central in the organization of TZ complexes, creates the same or different phenotype from MKS3 RNAi.

On the dorsal surface of the cell, all single basal body units are ciliated and, therefore have a transition zone. In the paired basal body unit, only the posterior basal bodies are ciliated and therefore will be the only basal bodies with a transition zone. With this in mind we fused a GFP tag to the N-terminus of B9D2, a known transition zone protein, in an effort to distinguish the posterior basal body of a pair from the anterior basal body. Distinguishing between the basal bodies of a pair would allow us to determine whether only the anterior basal bodies become misaligned from the straight rows of the basal bodies on the dorsal surface upon MKS3 or MKS5 depletion.

In this study, we show depletion of MKS5 does not affect the straight basal body rows from the posterior pole towards the anterior pole of the cell although some areas on the dorsal surface of the cell show ciliary loss. This study shows that GFP tagged *Paramecium* B9D2 localizes above the basal body, in an area consistent with transition zone (in single basal body unit), and localizes to only the posterior basal body in two basal body units. We attempted MKS3 RNAi on cells expressing GFP tagged *Paramecium* B9D2 in an effort to determine if the anterior basal body is the cause of the observed basal body misalignment.

The results of the study were inconclusive. GFP tagged B9D2 no longer discreetly localizes to the transition zone of the posterior basal body in MKS3 depleted cells. As a result we were unable to distinguish the anterior basal body from the posterior basal body in MKS3 depleted cells expressing GFP tagged *Paramecium* B9D2. Depletion of MKS 5 in GFP tagged *Paramecium* B9D2 expressing cells clearly demonstrates MKS5 protein plays an important role in the localization of B9D2 in the basal body unit in *Paramecium* although MKS5 does not play a role in the maintenance of the straight basal body rows.

Materials and methods:

Stocks, cultures and chemicals

Cells (stock 51s *P. tetraurelia*, sensitive to killer) were grown in wheat grass medium inoculated with *Aerobacter aerogenes* (Sasnar and Van Houten, 1989). All chemicals were purchased from Sigma-Aldrich (St Louis, MO, USA) unless otherwise noted.

RNAi constructs:

RNAi construct for *MKS5* gene (Genes ID number: GSPATG 00016244001) was designed from the sequences (genes nucleotide base positions +785 to +1436) in the *Paramecium* annotated genome using the ParameciumDB database (http://paramecium.cgm.cnrsgif.fr/cgi/alignment/off- target.cgi). Off-target analysis for the construct was done in ParameciumDB (http://paramecium.cgm.cnrsgif.fr/cgi/alignment/off- target.cgi). Specific oligonucleotide primers (forward primer: 5′GTAGGCAATGATGAACTCTAAG3′; reverse primer: 5′CTCCAAGCAGGATCTTCTC3′) were used to amplify the designed sequence using genomic DNA as a template. Genomic DNA was extracted by organic extraction as described in Picariello *et al*., 2014. PCR amplicons were then cloned directly into pCR2.1-TOPO vector (Invitrogen/Life Technologies), transformed into bacteria and sequenced according to the manufacturer (Invitrogen/Life Technologies) instructions. Correct insert sequences were cut from the pCR2.1-TOPO vector and ligated into the double T7-promoter vector L4440 (Addgene, Cambridge, MA, USA) using the quick ligation™ reaction kit (New England BioLab Inc, Ipswich, MA, USA) as per the kit

instructions. Vector insert was sequenced to confirm the presence of correct sequence. MKS5 RNAi plasmid construct was maintained in *Escherichia coli* DH5α bacteria at -80^{0}C in a 30% glycerol stock.

GFP-tag of B9D2 construct preparation and plasmid injection:

To localize B9D2, we added the coding sequence for GFP to the 5′ end of the genomic DNA sequence for in the pPXV plasmid using the restriction enzymes Nhe I and Xho I (New England Biolabs, Ipswich MA, USA). These cut sites were created using large primers to add them to either end of the sequence: forward (5′-gcggctagcatgctcgcagacagtcca -3′) and reverse (5′-gcgctcgagtcattaattcaaataatttctaatag-3′). We used GFP-pPXV plasmid (courtesy of Dr. W. John Haynes, University of Wisconsin, Madison, WI, USA) to clone B9D2 gene. The gene sequence is available in ParameciumDB (http://paramecium.cgm.cnrs-gif.fr/). Accession numbers: GSPATG00034796001. The target gene sequence was amplified by using Q5TM Hot Start high fidelity DNA polymerase (New England BioLab Inc, Ipswich, MA, USA) according to the manufacturer's protocol. Resulting amplicons were inserted into the pPXV plasmid using restriction enzymes (Nhe I and Xho I) (New England BioLab Inc, Ipswich, MA, USA) and the amplicon sequences were confirmed by sequencing. All the GFP- pPXV plasmids containing the target gene sequences were linearized by using Not I restriction enzyme (New England BioLab Inc, Ipswich, MA, USA). The linearized plasmid was purified and resuspended in water at a concentration of 5-10 µg/µl. Approximately 10-15 wild type 51s *P. tetraurelia* cells which had recently undergone autogamy were injected with 5 to 50 pg of the linearized plasmid into the macronucleus. Individual cells were transferred to 500 µl of culture fluid in depression slides after

injection and incubated at room temperature for 24hrs to recover and divide. Cells in each depression were then transferred to 10 ml culture fluid and maintained at 15^0C as individual clones. The presence of GFP tagged-B9D2 plasmid in individual clones was tested using PCR with extracted genomic DNA as a template. Genomic DNA was extracted from the clone cultures as described previously and tested with PCR using plasmid-specific primers: the forward primer was the same forward primer listed above and a reverse primer (5′-ttatttaagtgttgttcattta-3′), which was specific for the plasmid. One µl (approximately 400 ng) of DNA was used in each PCR reaction: one cycle of 95°C for 5 minutes followed by 30 cycles of 95°C for 1 minute, 44°C for 1 minute, and 68°C for 1 minute, followed by a final 5 minute extension step at 68°C. PCR products were analyzed on a 1% agarose gel and were stained with ethidium bromide.

RNAi feeding:

Escherichia coli strain HT115 (DE), which lacks RNaseIII, was used for RNAi feeding. HT 115 was transformed with 50ng of RNAi plasmid constructs of interest. HT115 bacteria transformed with L4440 with no insert were used as control. Overnight cultures of HT 115 bacteria transformed with a MKS5 RNAi construct or control were used to inoculate 50 ml of luria broth containing ampicillin (100 µg/ml) (LB-AMP) cultures. The bacterial cultures were incubated at 37^0C in a shaking incubator (New Brunswick scientific) until they reached 595 nm optical densities (OD) 0.3-0.4. At the desirable OD (0.3-0.4), cultures were induced with isopropylthio-β-galactoside (IPTG) (RPI Corp., Mt. Prospect, IL, USA) by adding to a final concentration of 125 µg/ml. Cultures were then incubated at 37^0C in a shaking incubator for additional 4hrs to produce double stranded RNA. The induced bacteria cultures were centrifuged at

3,439xg for 10 minutes at 4^0C (Beckman Coulter, Brea, CA, USA) and the bacterial pellets were resuspended in 100 ml of wheat grass media containing additional stigmasterol (8µg/ml), ampicillin (100µg/ml) and IPTG (125 µg/ml). For washing the paramecia cells, harvested cells (by centrifugation, 500xg) were resuspended in 100 ml Dryl's solution (1 mM Na_2HPO_4, 1 mM NaH_2PO_4, 1.5 mM $CaCl_2$, 2 mM Na-citrate; pH 6.8) and again harvested by centrifugation (500xg). *Paramecium* cells which had recently undergone autogamy (checked by Dippell staining) (Dippell, 1955) were washed in Dryl's solution (1 mM Na_2HPO_4, 1 mM NaH_2PO_4, 1.5 mM $CaCl_2$, 2 mM Na-citrate; pH 6.8) and approximately 200-300 paramecia cells added to the 100 ml wheat grass media containing the induced bacterial culture. All RNAi cultures were maintained at room temperature. An additional amount of induced bacteria (50 ml of induced bacteria culture), stigmasterol (8µg/ml), ampicillin (100µg/ml) and IPTG (125 µg/ml) were added to the culture at 24hrs and 48hrs after feeding. Paramecia cells collected by centrifugation (500xg) at 72hrs after feeding and subjected to further analysis (immunofluorescence and the extraction of m-RNA) unless noted otherwise. The RNAi for meckelin (MKS3) was done by following the procedure described in Picariello et al., 2014. All the RNAi experiments were repeated a minimum of three times.

To investigate how the depletion of MKS5 affects the location of B9D2 protein we used paramecia cells injected with the GFP-B9D2 construct. We followed the procedure described above for the depletion of meckelin (MKS3) in GFP-B9D2 construct injected cells.

Reverse transcription-PCR (RT-PCR)

We used RT-PCR to check the efficacy of the RNAi feeding procedure by following the protocol published by Yano et al., 2003. RT-PCR was used to check the efficacy of the RNAi feeding. We consider the data obtained from the RT-PCR to be semi-quantitative. A pair of calmodulin primes (forward primer: 5′-CTGAAGCTGAACTTCAAG-3′; reverse primer: 5′-CAGAATGATGGTTTCTAAATGA-3′) was used in RT-PCR as a quality control for method. Three different concentrations (undiluted, diluted 10 fold and 100 fold) of cDNA were used as template. All the experiments were repeated a minimum of three times.

Immunofluorescence:

Paramecia cells after RNAi feeding (100 ml culture) were harvested by centrifugation (500xg) (Damon IEC Division Clinical Centrifuge, Needham Heights, MA, USA). Cells were washed three times in 100 ml Dryl's solution by resuspending the cells in Dryl's solution followed by harvesting the cells by centrifugation (500xg). Finally, harvested cells were resuspended in ~100µl of Dryl's solution. Cells were permeabilized in PHEM buffer (60 mM piperazine ethanesulfonic acid(PIPES), 25 mM hydroxyethyl piperazineethanesulfonic acid (HEPES), 10 mM ethylene glycol tetraacetic acid (EGTA), 2 mM $MgCl_2$ and 0.1% Triton X-100; pH 6.9) and fixed for 30-40 min with fixation buffer (2% or 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA, USA), 2 mM $NaH_2PO_4 \cdot H_2O$, 8 mM Na_2HPO_4, 150 mM NaCl; pH 7.5). Cells were washed three times in blocking buffer (2 mM $NaH_2PO_4 \cdot H_2O$, 8 mM Na_2HPO_4, 150 mM NaCl, 10 mM EGTA, 2 mM $MgCl_2$, 1% Tween20, 1% BSA; pH 7.4) by resuspending the cells in blocking buffer followed by harvesting the cells by

centrifugation (500xg). Paramecia cells were immunostained by incubating at room temperature for 1 hr in primary antibody solution (100µl). All antibodies were diluted in blocking buffer.

For the visualization of the SRs and basal bodies, primary antibodies were as follows: rabbit anti-SR at a dilution of 1:400 (gift from Janine Beisson, Centre de Génétique Moléculaire, Gif-sur-Yvette, France) and mouse anti-Glu-α-tubulin at a dilution of 1:500 (Synaptic Systems, Göttingen, Germany). For the visualization of cilia, primary antibody was as follows: mouse anti-α-tubulin at a dilution of 1:200 (Sigma-Aldrich, St Louis, MO, USA). Cells were washed three times in washing buffer (2 mM $NaH_2PO_4 \bullet H_2O$, 8 mMNa_2HPO_4, 150 mM NaCl, 1% Tween20; pH 7.4) after primary antibody incubation by resuspending the cells in washing buffer followed by harvesting the cells by centrifugation (500xg).Then cells were incubated in secondary antibody at room temperature for 1hr. Secondary antibodies were Alexa Fluor 488 goat anti-mouse, Alexa Fluor 555 goat anti-mouse, Alexa Fluor 488 goat anti-rabbit, or Alexa Fluor 568 goat anti-rabbit (Molecular Probes/Invitrogen, Grand Island, NY, USA). After incubation in secondary antibodies, cells were washed three times with washing buffer (2 mM $NaH_2PO_4 \bullet H_2O$, 8 mMNa_2HPO_4, 150 mM NaCl, 1% Tween20; pH 7.4) by resuspending the cells in washing buffer followed by harvesting the cells by centrifugation (500xg). Finally, cells were suspended in Vectashield® mounting medium (Vector Labs, Burlingame, CA, USA) for imaging. All the images were captured using the DeltaVision® restoration microscopy system (Applied Precision), consisting of an inverted Olympus IX70 microscope (Olympus America, Center Valley, PA, USA) and a Kodak CH350E camera (Rochester, NY, USA).

For the visualization of the B9D2 protein along with basal bodies in GFP-B9D2 construct injected paramecia cells the primary antibodies were as follows: mouse anti-GFP at a dilution of 1:200 (Sigma-Aldrich, St Louis, MO, USA) and rabbit anti-centrin at a dilution of 1:1000 (anti-Tetrahymena centrin, gift from Mark Winey, University of Colorado, Boulder, CO, USA).

Results:

Bioinformatics analysis and RNAi construct:

The MKS5 (also known as RPGRIP1) protein from *Homo sapiens* (Accession number: NP_001295263.1) was used for searching the MKS5 protein sequences in *Paramecium* annotated genome in the dedicated database ParameciumDB (http://paramecium.cgm.cnrs-gif.fr/). We identified two *Paramecium tetraurelia MKS5* genes (GSPATG00016244001; GSPATG00012529001) that encode MKS5 proteins using the *H. sapiens* MKS5 protein (1269 amino acids long) in BLAST search. The MKS5 protein in *Paramecium* (839 amino acid long) has four coiled coil domain with one C2 domain where as human MKS protein has five coiled coil domain with two C2 domain.

RNAi construct for *MKS5* gene (GSPATG0001624400; nucleotide position: +785 to +1436) was created from the sequences in the *Paramecium* annotated genome using the ParameciumDB database. Off-target analysis for the construct was done in ParameciumDB (http://paramecium.cgm.cnrsgif.fr/cgi/alignment/off- target.cgi). Our RNAi off target analysis predicted MKS5 RNAi construct would down regulate both copies *MKS5* genes in *Paramecium*.

Depletion of MKS5 leads to loss of cilia but does not affect the straight basal body rows:

We depleted MKS5 transcripts in *Paramecium* by RNAi. Depletion of message was checked by RT-PCR using specific primers for *MKS5* gene. Agarose gel analysis of

the RT-PCR clearly demonstrates the depletion of message in the MKS5 RNAi cells compared to control L4440 fed cells (Fig.3.1).

Both the control and MKS5 RNAi cells were treated with mouse-α-tubulin to visualize cilia on the cell surface (Fig.3.2). In the control cell, both the dorsal and ventral surfaces of the cells are covered with numbers of cilia (green) (Fig.3.2A). In the MKS5 mRNA depleted cells, dorsal surface of the cells have fewer cilia compared to the control cells (Fig. 3.2B). Some areas of the cell surface showed complete baldness (loss of cilia). We observed the same phenotype in the ventral surface of the cell. Loss of cilia was also evident in the MKS3 protein depleted *Paramecium* cell in previous study (Picariello *et al.*, 2014).

The basal body rows along with SRs in the L4440 fed control cells, MKS3 RNAi cells and MKS5 RNAi cells were examined using immunofluorescence. In all cells, anti-1D5 labels the basal bodies and anti-SR labels the SRs. The whole cell images are comprised of the stack of Z-sections approximately 10 μm thick to ensure all basal bodies along with SRs are visible. In the control cells (Fig.3.3A), basal bodies (green) and SRs (red) have a highly organized characteristic pattern. Basal bodies remain in straight rows from the posterior pole to the anterior pole both on the dorsal and the ventral surfaces of the cell. SRs originate from the basal bodies and extend toward the anterior pole of the cell and traverse several anterior basal body units (Fig.3.3B). In the MKS5 transcript depleted cells we observed the same phenotype both in the dorsal and ventral surfaces compare to the control cells (Fig.3.3C). The basal body rows remain straight along with a highly ordered organization of the SRs from pole to pole of the cell (Fig.3.3D). In the MKS3 transcript depleted cells the basal body row alignment and the SRs appearance are

dramatically different (Fig.3.3E). The basal body rows appear convoluted and the SRs have a disordered organization (Fig.3.3F). Misaligned basal bodies along with the SRs are observed predominantly on the middle of the dorsal surface of the cells.

Depletion of MKS5 affects the localization of the B9D2 protein in the transition zone (TZ) of the basal body:

To investigate how MKS5 or MKS3 transcript depletion affects the localization of B9D2 protein (a known marker for TZ of basal body) we used *Paramecium* cells expressing GFP-B9D2 protein for RNAi experiments followed by immunofluorescence. All the cells (control L4440 fed cells, MKS5 depleted cells or MKS3 depleted cells) were treated with anti-centrin and anti-GFP antibody to stain basal body (red) and GFP-B9D2 (green) protein respectively. In the control RNAi cells, basal body rows remain straight from the posterior pole towards the anterior pole of the cell. Immunofluorescence signal (green) for GFP-B9D2 showed some diffuse intracellular localization but predominantly concentrated above the plane of the basal body which is consistent with its expected localization in the TZ (Fig.3.4A). Furthermore, the localization of GFP-B9D2 protein allowed us to distinguish the anterior basal body from the posterior basal body in two basal body units (Fig.3.4B). In the MKS5 protein depleted cells, like the control cell, basal body rows remain straight from the posterior pole towards the anterior pole of the cell (Fig.3.4C). Anti- GFP-B9D2 protein staining clearly demonstrates depletion of MKS5 affects the localization of GFP-B9D2 protein to the anterior basal body of the two basal body unit (Fig.3.4D). Moreover, localization of the GFP-B9D2 was also disrupted for the single basal body unit since most of basal body lacks the signals for GFP-B9D2 (Fig.3.4D). In the MKS3 protein depleted cells, basal body polarity was disrupted

(Fig.3.4E) and GFP-B9D2 localization became dispersed such that GFP-B9D2 no longer discreetly localized to the TZ of posterior basal bodies and can be found throughout the cell membrane. We were unable to track the anterior basal body of the pair and show that it is responsible for the loss of the basal body polarity and cortical ridge distortions (Fig.3.4F).

We also used GFP-empty vector expressing *Paramecium* cell to perform the RNAi for the *MKS5* gene. The L4440 fed GFP-empty vector expressing cell served as a control. Both the cells were treated with anti-centrin and anti-GFP antibody to visualize the basal body unit (red) and the GFP protein respectively. In the control cell, basal body (red) rows remain straight from the posterior pole toward the anterior pole of the cell. There was almost no signal for the GFP protein (green) (Fig.3.5A). In the MKS5 transcript depleted cell we also observed the straight basal body rows like the control cell. In these cells there was also almost no signal for the GFP protein (green) (Fig.3.5B).

Discussion:

In this study we demonstrate the function of MKS5 protein in *Paramecium*. It is a follow up study of the observation that has reported from our lab previously (Picariello et al., 2014). In this study (Picariello *et al*., 2014), authors reported depletion of TZ protein MKS3 transcripts by RNAi cause loss of cilia in *Paramecium*. The striking observation upon the silencing of MKS3 was the misaligned basal body rows along with a disordered organization of the SRs. MKS3 is the part of MKS/B9 protein complex that comprise the ciliary basal body transition zone (TZ) that forms a molecular gate to control the entry and exit of various protein molecules inside ciliary axoneme. Here in this study, we demonstrated the role of MKS5 protein, another TZ protein in the maintenance of basal body rows from the posterior pole towards the anterior pole, the ordered organization of basal body associated SR and the biogenesis of cilia in *Paramecium*.

We used RNA interference to silence the transcript of MKS5 in *Paramecium*. In *Paramecium* there are two genes (two genes are ohnologs of each other: GSPATG00016244001 and GSPATG00012529001) that code for MKS5 protein. Our off target analysis of the RNAi construct (we used the gene sequence of GSPATG00016244001 to create our construct) demonstrated the silencing of both the transcripts of MKS5. We also validate the depletion of m-RNA for the MKS5 by doing RT-PCR which clearly demonstrates the depletion of MKS5 transcripts in the test cell. Our immunofluorescence data of cilia staining (using α-tubulin) in the L4440 fed control cell and MKS5 transcripts depleted test cell clearly demonstrates silencing of MKS5 transcripts results in loss of cilia both on the dorsal and ventral surface of the cell. This is an expected result since data in other system describes MKS5 which is a part of NPHP1-

4 protein complex of the TZ, plays an important role in interacting with other MKS module proteins to form the TZ of the cilia (Czarnecki and Shah, 2012). Loss of cilia in MKS5 silencing *Paramecium* cell is also in accordance with MKS3 depletion in *Paramecium* cell (Picariello et al., 2014). The data obtained from this study clearly suggests MKS5 protein in *Paramecium* plays a role in the biogenesis or the maintenance of the cilia.

We demonstrated the role of MKS5 in maintaining the basal body row alignment and the ordered organization of the SR on the cell surface by immunofluorescence. To examine the role of MKS5 we treated the control, MKS3 depleted cells or the MKS5 depleted cells to mark the basal boy units and the SRs by anti-1D5 and anti-SR antibody respectively. Unlike MKS3 depleted cell, in MKS5 depleted cell basal body row remains straight from the posterior pole towards the anterior pole of the cell. The basal body associated SRs organization remains ordered in the MKS5 transcripts depleted cell. This phenotype of the MKS5 transcripts depleted cell is similar to the phenotype of the control L4440 fed cells and in contrast to the MKS3 transcripts depleted cell where we observed misaligned basal body rows along with a disordered organization of the SRs predominantly in the dorsal mid region of the *Paramecium* cell. All the data suggests although MKS5 depletion in *Paramecium* causes ciliary loss but does not affect the highly organized basal body rows and associated SRs.

The method of using GFP-B9D2 expressing cell for silencing the MKS5 and MKS3 transcripts allowed us to demonstrate the role of MKS5 or MKS3 in localizing another TZ protein B9D2. B9D2 protein is a part of MKS/B9 multiprotein complex that is found in TZ of the basal body. The data obtained from the study clearly suggests

silencing of MKS5 or MKS3 affects the localization of another transition zone component, B9D2. The mis-localization/ absence of signals for GFP-B9D2 (in the basal body units) in the MKS5 silenced cells clearly suggest the role of MKS5 protein, the member of NPHP 1-4 multiprotein complex may affect the localization of MKS/B9 multi-protein complex. The mis-localization or loss of GFP-B9D2 (in the basal body units) signals in MKS3 silenced cells suggests that at least one MKS/B9 multiprotein complex member cannot be incorporated or maintained in the TZ. These data support the idea that the TZ becomes compromised or silenced in response to the silencing of MKS5 or MKS3 in *Paramecium.*

It is intriguing to think that NPHP1-4 multi-protein complex and MKS/B9 multi-protein complex may interact with one another to form or maintain the TZ in *Paramecium*. In the absence of any member of the above mentioned multi-protein complex the TZ become destabilized. This speculation explains the global lack of cilia and the altered localization of a conserved transition zone component (B9D2) in MKS5 or MKS3 silenced cells. It would be interesting to determine the composition of the *Paramecium* TZ complex e.g., NPHP1-4 complex, MKS/B9 complex, NPHP 5-6 complex or the member of inversion compartment complex. Another interesting aspect would be to know the function of each of the protein that is the part of different multi-protein complex. Relatively little is known about the composition of the TZ in *Paramecium*. Our work gives an idea how to explore the function of different proteins that form the TZ in *Paramecium tetraurelia*. Determining the proteins present in this region could serve two purposes. First it would define the protein compliment of the

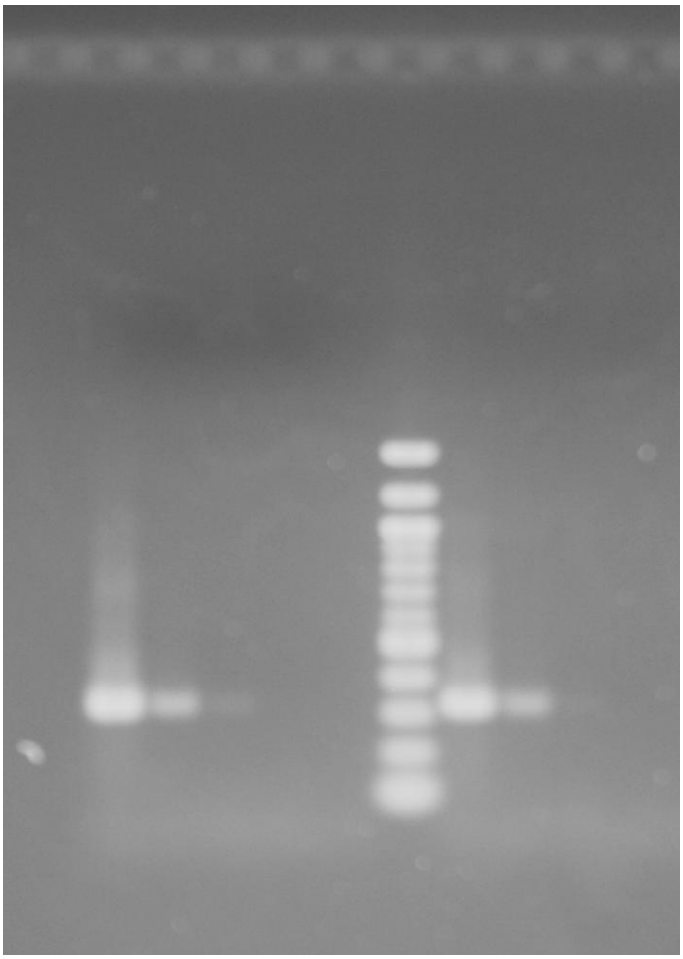
Paramecium TZ. Second it might provide a way to explore the any noble role of TZ protein outside their function in TZ of the basal body.

Figures:

Fig. 3.1: RT-PCR analysis showing the depletion of *MKS5 gene* specific m-RNA in test cells compared to control cells.

Panel A shows a representative image of an agarose gel demonstrating the level of mRNA for *calmodulin* gene using *calmodulin* gene specific primers in control L4440 fed cell or test (*MKS5* gene depleted) cells. In Panel A, m indicates the 100 bp DNA Ladder and a, b and c indicates 1:10, 1:100 and 1:1000 dilution of cDNA which was made using reverse transcriptase enzyme after extraction of total mRNA from the control cell or the test cell. The d indicates 1:10 dilution of cDNA which was made without using reverse transcriptase enzyme after extraction of total mRNA from the control cell or the test cell.

Panel B shows a representative image of an agarose gel demonstrating the level of mRNA for *MKS5* gene using *MKS5* gene specific primer in control L4440 fed cell or test (*MKS5* gene depleted) cells. In Panel B, m indicates the 100 bp DNA Ladder and a, b, c and d indicates 1:10, 1:100, 1:200 and 1:1000 dilution of cDNA which was made using reverse transcriptase enzyme after extraction of total m-RNA from the control cell or the test cell. The e indicates 1:10 dilution of cDNA which was made without using reverse transcriptase enzyme after extraction of total mRNA from the control cell or the test cell.


Control cell
MKS5 RNAi cell
Calmodulin gene
a b c d m a b c d cDNA dilutions

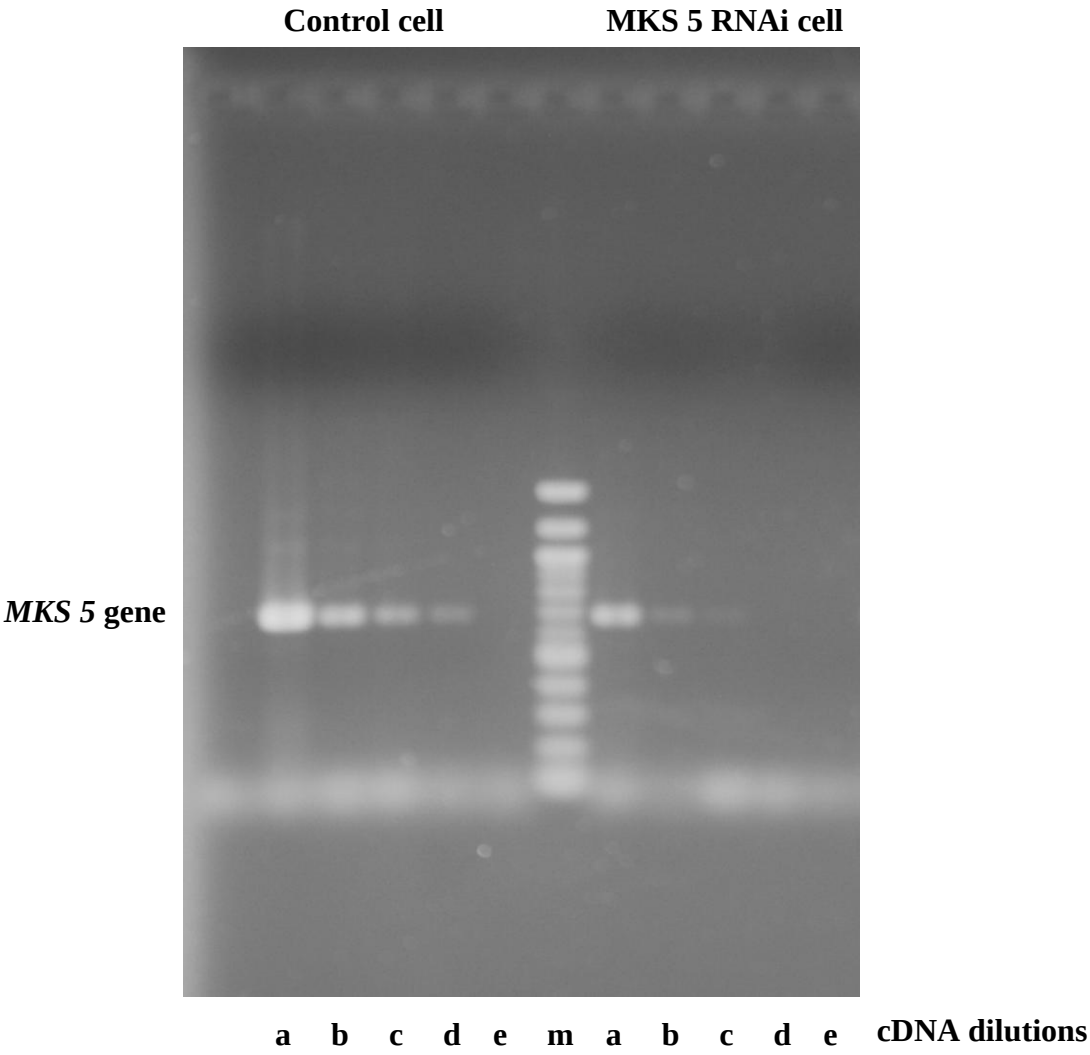
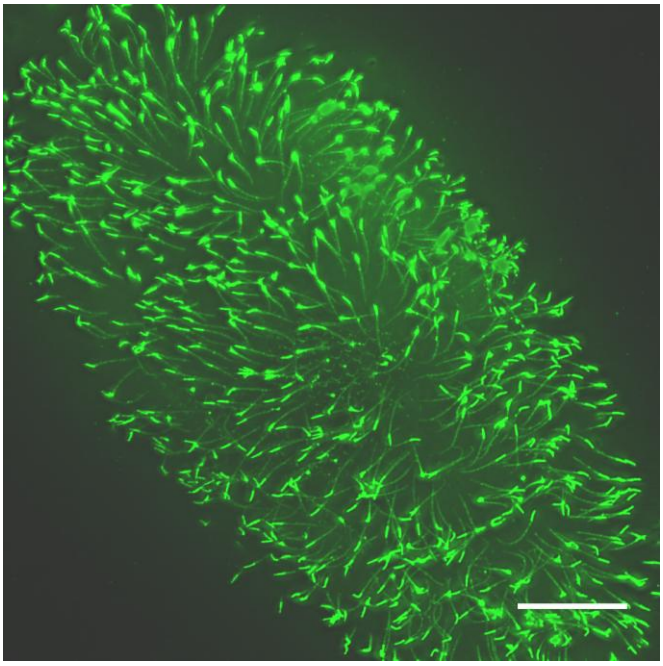
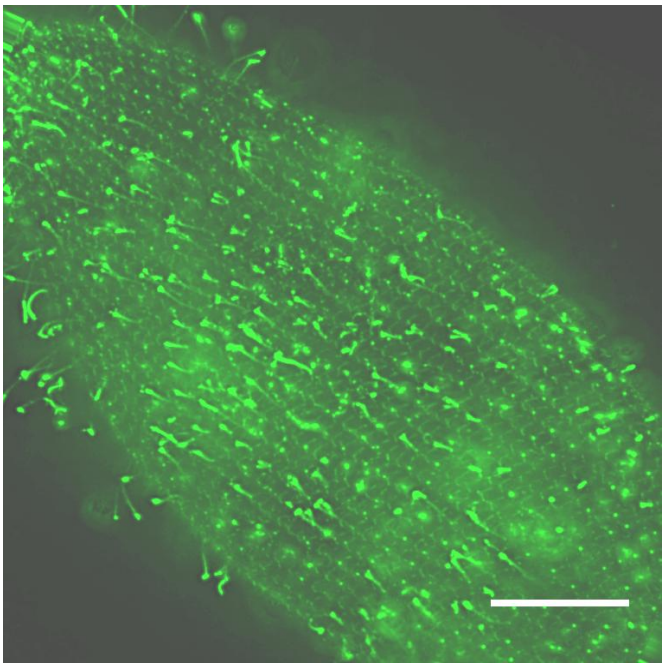

Control cell
MKS 5 RNAi cell
MKS 5 gene
a b c d e m a b c d e
cDNA dilutions

Fig. 3.2: Depletion of *MKS5* gene leads to loss of cilia on the cell surface.

All cells were treated with α-tubulin antibody to visualize the cilia. In Panel A, cell surface of the control L4440 fed cell is covered with number of cilia (green). In Panel B, in the test cell silencing of MKS5 transcript leads to ciliary loss. We observe loss of cilia both on the dorsal and the ventral surface. Scale bar is 15μm.



Panel A: Cilia (green) in control cell

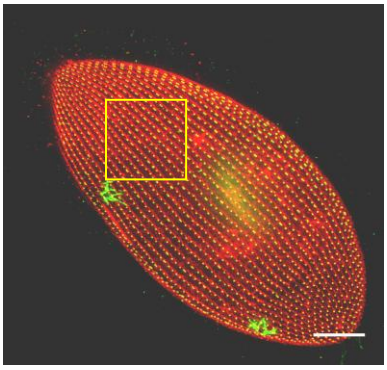


Panel B: Cilia (green) in test (MKS5 depleted cell)

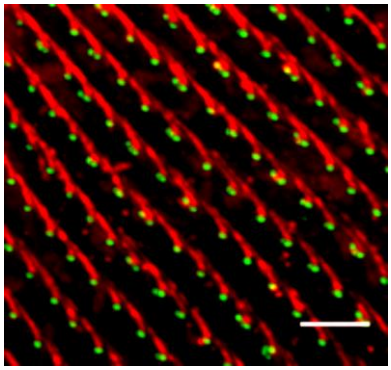
Fig. 3.3: Depletion of MKS5 cause loss of cilia but does not affect the straight basal body rows and the ordered organization of the SRs.

All the cells were treated with anti-1D5 and anti-SR antibody to visualize the basal body units (green) and SRs (red) respectively on the cell surface. Panel A and B show the basal body row alignment and SRs organization in the control cell. In the control cell basal body rows remain straight and SRs show a highly ordered organization. SR emanates from the basal body unit and extends towards the anterior pole of the cell. Panel C and D show the basal body row alignment and SRs in the MKS5 depleted cells. The phenotype of the cell is similar to the control cell. Panel E and F show the basal body row alignment and SRs in the MKS3 depleted cell. In the MKS3 depleted cell, basal body rows are misaligned and SRs have a disordered organization on the cell surface. Scale bars are 15μm in A, C or E and 3μm in B, D or F.

Control cell

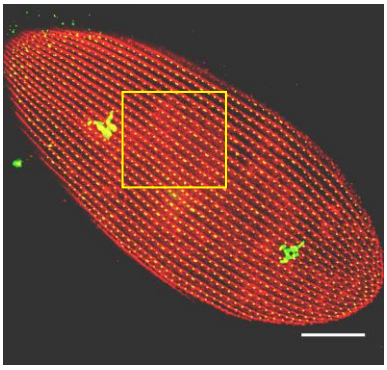


A

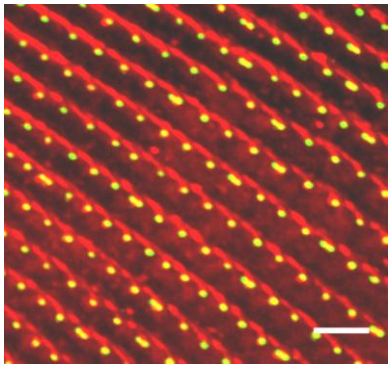


B

MKS5 depleted cell

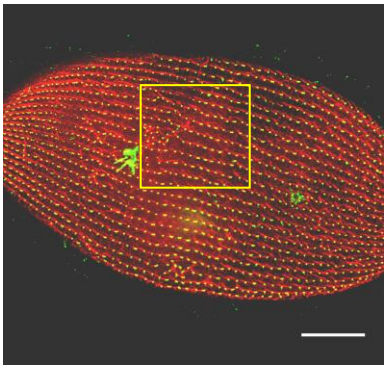


C

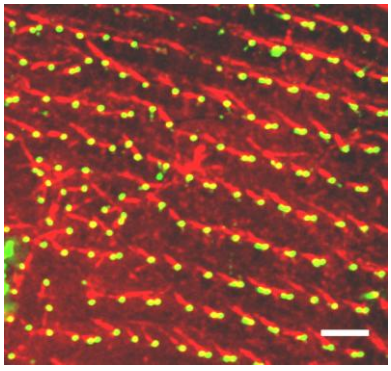


D

MKS3 depleted cell



E

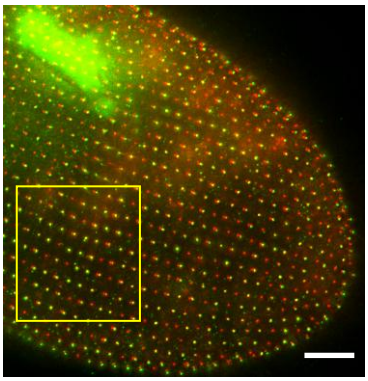


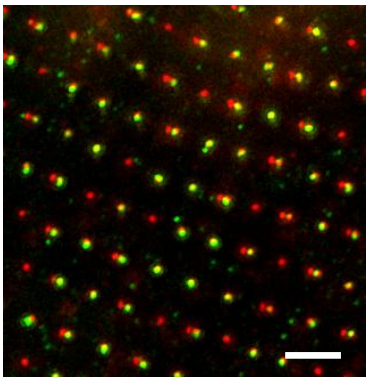
F

Fig. 3.4: Depletion of MKS5 affects the localization of B9D2 protein in TZ of the basal body.

All the cells were treated with anti-centrin and anti-GFP antibody to visualize the basal body units (red) and GFP-B9D2 protein in the cell. In all the images, the yellow box highlighted area is enlarged to show the basal body rows and GFP localization in the GFP-B9D2 expressing cell. Panel A and B show the basal body row alignment and GFP-B9D2 localization in the control cell. In the control cell, basal body (red) rows remain straight with GFP-B9D2 protein localization at the TZ of the basal body in a single basal body unit. In two basal body unit only the posterior basal body has the signal for GFP-B9D2 protein. Panel C and D show the basal body row alignment and GFP-B9D2 protein localization in the MKS5 depleted cell. In the MKS5 depleted cell, basal body rows remain straight like the control cell but GFP-B9D2 localization shows remarkable difference compared to control cell. Both the single basal body and the posterior basal body of the two basal body unit lack the localization of B9D2 protein in the MKS5 depleted cell. Panel E and F show the basal body row misalignment and localization of B9D2 protein in the MKS3 depleted cell. In the MKS3 depleted, cell basal body rows are misaligned and GFP-B9D2 protein shows very dispersed and diffused localization. Scale bars are 10μm in A, C or E and 3μm in B, D or F.

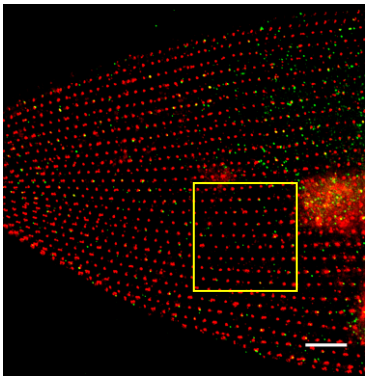
Control cell

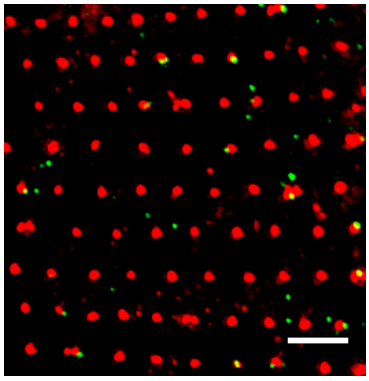




A **B**

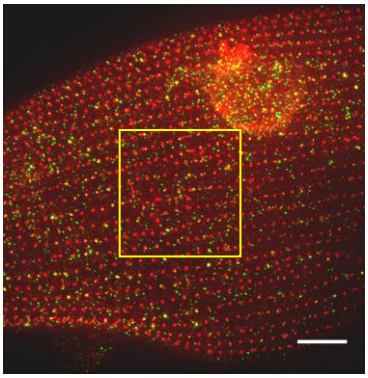
MKS5 depleted cell

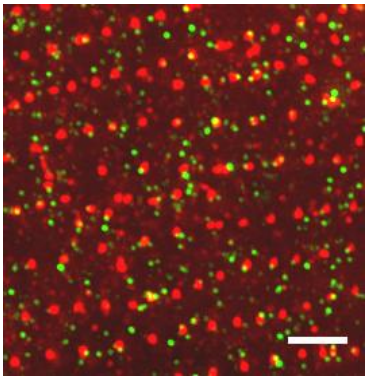




C **D**

MKS3 depleted cell

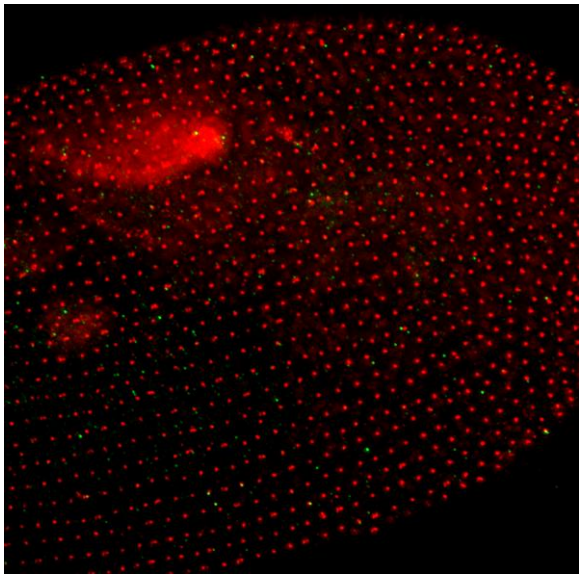




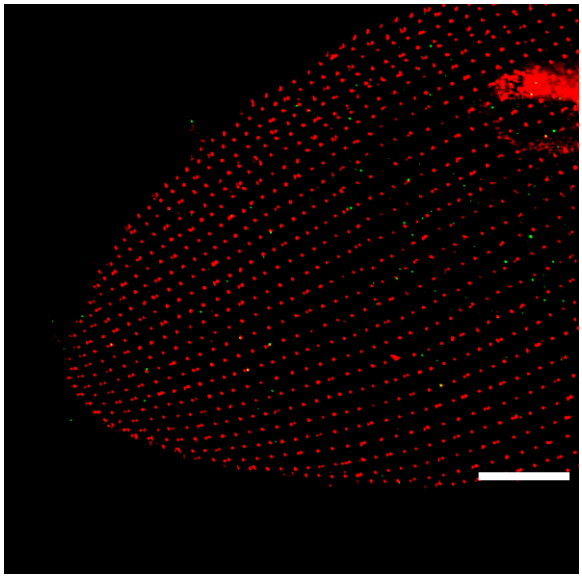
E **F**

Fig. 3.5: Depletion of MKS5 in GFP-empty vector expressing cells.

All the cells were treated with anti-centrin and anti-GFP antibody to visualize the basal body units (red) and GFP-B9D2 protein in the cell. Both the control and MKS5 depleted cells were injected with GFP-empty vector. Panel A shows the basal body row alignment in the control cell. In the control cell, basal body (red) rows remain straight from the posterior pole toward the anterior pole of the cell. There is almost no signal for the anti-GFP antibody (green). Panel B shows the basal body row alignment in the MKS5 depleted cell. In the MKS5 depleted cell, basal body rows remain straight like the control cell. MKS5 depleted cell also do not show any signal for the anti-GFP antibody (green). Scale bars are 10 μm.



A (Control cell)



B (MKS 5 depleted cell)

Chapter 4: Concluding remarks

Conclusions

In this study, we demonstrated that a number of SR proteins in *Paramecium* (at least one from each of five Structural Groups) must be concurrently expressed to form an SR of appropriate length and striation. We do not know the location of each protein relative to the others in the SR, but we assume that there must be physical interactions, perhaps through their coiled coil domains to create a striated structure of the normal length.

The SR proteins in *Paramecium* are categorized into five Structural Groups based on their ability to form the coiled coil domains. The functional study (by RNAi of the SRs) data clearly demonstrates that members of all five Structural Groups are required to form an SR of appropriate length and striation. The data presented here support a reciprocal relationship between SR and the Infraciliary Lattice (ICL, inner most cytoskeleton structure of the *Paramecium* cortex). Depletion of individual SR Structural Groups leads to the abnormal appearance of SRs and causes basal body misalignment. In addition to this, depletion of SR Structural Groups also suggests that SRs play an important role in shaping the morphology of the cell surface or other cytoskeleton structures (e.g., cortical units, epiplasm and ICL) possibly by maintaining the shape of the ICL mesh unit. There seems to be a reciprocal relationship of the SR with the ICL, in that disruption of the ICL affects the orientation of the SRs. These findings have shown a previously unknown function of the SR.

The SR is a cytoskeletal structure, present in different systems such as ciliates (Rubin and Cunningham, 1973), green flagellates (Lechtreck and Melkonian, 1991;

Weber et al., 1993), ciliated epithelial cells (Stephens, 1975; Lemullois et al., 1991) and mammalian cells (Yang et al., 2002; Yang et al., 2005). The SR in *Chlamydomonas,* one of the best studied SR structures, is made of only a single polypeptide, known as SF-assemblin, a 34 kDa protein (284 amino acid residues). The SF-assemblin protein has two domains: one is the N-terminal short head domain (31 amino acid residues) and other is the C-terminal helical rod domain, which has the ability to form a coiled coil structure (Lechtreck and Melkonian, 1991; Weber et al., 1993). In mammalian systems, the SR is the prominent structure found at the proximal end of the basal body that extends towards the nucleus. As in *Chlamydomonas,* in the mammalian system, the SR is known to comprise only one protein, rootletin, a 220 kDa protein (2017 amino acid residues). The rootletin protein has an N-terminal short head domain and a C-terminal rod domain, which has the ability to form coiled coil structure (Yang et al., 2002; Yang et al., 2005).

In ciliates, unlike *Chlamydomonas* and mammalian systems, there are many SR genes. We asked the question: are all of them required to make a functional full length SR in the cell? Our blast search approach demonstrates 30 potential SR proteins in *Paramecium*. All of these sequences code for SR proteins that have the characteristic domains of an SF-assemblin protein. The phylogenetic analysis of these 30 proteins categorizes them in 13 Paralog Groups. Localization studies of representative members of 13 paralog group using N-FLAG tagging and immunofluorescence demonstrate that those SR proteins that have the domains of SF-assemblin localize throughout the length of SR. We combined all 13 SR Paralog Groups into 5 Structural Groups based on their potential to form coiled coil domains in the same part of the proteins.

In an effort to understand how these Structural Groups are important to form a full length SR we utilized RNAi to deplete the transcripts of the Structural Groups. RNAi data of Structural Group depletions clearly demonstrate that at least one member from each Structural Group is must be expressed to form an SR of full length and striation pattern. Together, these data have yielded a new insight into the importance of SR Structural Group members and their coiled coil motifs in the formation of the highly complex SR structure in ciliates.

In an effort to understand how proteins of these 5 Structural Groups interact with each other, we utilized the knowledge of the previously published models on the SF-assemblin and rootletin proteins (Lechtreck and Melkonian, 1998; Yang et al., 2002). In both systems, the individual protein monomers (SF-assemblin or rootletin) interact with each other to form a homodimer. The homodimers interact with each other to form a rootlet of the normal length and, where they overlap laterally to form a protofilament structure with the characteristic striations of the SR. The coiled coil structure of the rod domain plays an important role in the formation of the SR in both systems primarily through ionic interactions. The predicted length of SF-assemblin rod domain is ~36nm. The SF-assemblin protein molecules interact through the overlapping of the N-terminal and C terminal of the rod domain to establish the periodicity of ~28nm in the assembled SR in the *Chlamydomonas* (Lechtreck and Melkonian, 1998).

Based on the knowledge of the *Chlamydomonas* model and in combination with our functional studies data (RNAi of SR Structural Groups) and ultra structure studies data (striation pattern in isolated SRs from control cells and Paralog Group depleted cells), we propose a model (Fig.4.1) of how different Structural Group proteins in

Paramecium interact to form a full length SR with striations. The coiled coil structure analysis of the 5 Structural Groups in *Paramecium* demonstrates that, within a Structural Group, the SR protein members have very similar amino acid sequences (supplement Fig.S1). Moreover, the length and position of the coiled coil domains are remarkably similar within the members of a Structural Group. Our functional studies data provides the clue that, within a Structural Group, the loss of one SR member can be adjusted by other members of the same Structural Group. Therefore, we speculate the protein monomers of 5 different SR Structural Groups interact with each other longitudinally to form a pentamer. The ionic interaction of the coiled coil structures among the different SR protein molecules of the Structural Groups may play an important role in the creation of the pentamers, which we propose is a unit for building the SR. In the pentamer, at least one member of each of the 5 Structural Groups should present. RNAi studies of the different Structural Groups and Paralog Groups strongly support the idea of a pentamer, e.g., in the Structural Group 1 (Paralog Group 1: *SR1a* and *SR1b*) depletion of the transcripts of *SR1a* or *SR1b* separately does not show the signature phenotype of the SR Structural Group depletion and also does not have any effect on the length and striation pattern of the SRs. These data also suggest that individual SR proteins within a Structural Group may replace the function of other SR proteins within in the same Structural Group because of their ability to form the same coiled coil structure. The pentamers of the 5 Structural Groups overlap to form the longitudinal protofilament and also interact laterally to create more than one less dense regions that appear as striations (Fig.4.1). The full length SR forms by the bundling of protofilaments. The bundling arrangement of protofilaments results in the formation of two areas of high density (due to protein

packing densities) per periodicity. The SR Structural Group 1, 2, 3, 4 and 5 coiled coil rod domains are roughly ~34nm, 31nm, 32nm, 30nm and 31nm long. In this model, we are treating the proteins as rods, which may not be realistic. However, the length of the proteins as rods matches well with the striation periodicity. The bundling arrangements of protofilaments in our model could result in the formation of major and minor striations of 24-34nm periodicity.

While our model proposes a simplified method of how pentamers of proteins from the SR Structural Groups interact with each other to assemble SR, there are some other aspects of the model that remain unclear. The pentamers of the each Structural Group might organize in either parallel or anti-parallel fashion. The variation in striation periodicity and length of the SRs may occur due to the bundling of different numbers of protofilaments in the assembled SRs. There may be some regulatory proteins that may control the organization of the protofilaments to give rise to variable the length of the SRs. The SR is not of uniform width, and seems to be pointed at the anterior end. We do not know how the monomers of 5 different SR Structural Groups organize in the pentamer. Our immunofluorescence data of N-FLAG tagging of different SR protein molecules (resolution limit ~200nm) do not provide any insight into the organization. Although our ultra structure data of isolated SRs demonstrate the periodicities of major striation are 24-34nm, they do not provide any conclusive evidence how different members of SR Structural Group interact in the organization of the pentamer.

To get a clearer picture of how different SR Structural Groups interact with each other, we could carry out immunogold pair wise labeling of the different Structural Groups. Two different SR protein members (e.g., Structural Group 1 and Structural

Group 2) of two different Structural Groups (e.g., SR1a of Structural Group 1 and SR7a of Structural Group 2) could be tagged with different tags e.g., N-FLAG or N-MYC. The *Paramecium* cell co-injected with both the tagged in expression vectors would then be used for isolation of SRs, and the ST subjected to immunogold labeling (resolution limit: 5-10nm) using two different sized gold particles for the N-FLAG or N-MYC tags. The distance between two different sized gold particles should provide an idea of interactions between these two Structural Groups. This approach may provide the more concrete idea of interaction among the proposed pentamer proteins of Structural Groups.

In an effort to determine how the SR plays a role in maintaining the straight basal body rows, we utilized RNAi of SR Structural Groups followed by immunofluorescence for the basal bodies. Our immunofluorescence clearly demonstrates that SRs function in maintaining the straight basal body rows from the anterior pole toward the posterior pole on the cell surface. Even with basal body row misalignment, the angles between the rootlets themselves remain constant when compared to controls.

To understand a relationship between the SR structure and other cytoskeleton structures we analyzed the LC-MS/MS data from the isolated SRs. We found a clue of a relationship in between SRs and ICL1e subfamily protein of the ICL network since ICL1e subfamily proteins consistently co-isolated with SR proteins. Other workers showed that the LC-MS/MS data of the isolated ICL mesh co-isolate with SR proteins (25 SR proteins of 30 SR proteins) (Gogendeau et al., 2008). Moreover, the our LC-MS/MS data showed that the proteins of the ICL1e subfamily did not co-isolate with the SRs from cells depleted of Structural Groups (Structural Group 1 or Structural Group 2) With that in mind, we silenced the ICL1e subfamily members using RNAi. We also

silenced two other ICL subfamilies (ICL9 and ICL10), which other workers (Gogendeau et al., 2008) showed that did not lead to disruption of the ICL upon depletion by RNAi. In our studies, RNAi depletion of ICL subfamilies ICL9, ICL10 or ICL1e showed that that only the ICL1e subfamily plays an important role in the maintaining the straight basal body rows along with the ordered organization of the SRs. Data obtained from the SR Structural Group and ICL1e RNAi strongly suggest that there is a reciprocal relationship between these two cytoskeleton structures Fig.4.2 shows how the SRs interacts with the contractile cytoskeleton structure ICL mesh network through the ICL1e subfamily proteins to shape the entire cortex structure in *Paramecium*.

Both the SR structure and the ICL mesh network function together to maintain the straight basal body rows from the posterior pole towards the anterior pole of the cell. The interaction locations of these two cytoskeleton structures (ICL1e-SR) seem to be at a corner of the ICL mesh unit where the basal body attaches and at a side of the mesh unit 90 degrees away (see Figure 4.2) based on our immunofluorescence data and consistent with Beisson and workers' analysis of the ICL (Gogendeau et al., 2008). We propose that the SR holds the ICL mesh unit shape. If the length of the SRs is short due to loss of an SR Structural Group member these SRs cannot maintain the interaction between SRs and ICL and the shape of the mesh unit is distorted. Conversely, if ICL1e is not present (shown by the depletion of ICL1e transcripts), the interaction between SRs and ICL also cannot be maintained. Both the cytoskeleton structure also play a vital role in maintaining the shape of the other cytoskeleton structures e.g., cortical unit. We have established that if the SR or ICL is disrupted the other of these large cytoskeletal structures is likewise disrupted. The mis-shapen ICL mesh units in turn lead to the mis-shapen cortical units

that we see as ridges out of alignment and the epiplasm under the cortical unit likewise out of alignment. In other words, the changes at the level of the cortical unit (ridges and epiplasm) are a secondary effect of the disruption f the ICL mesh. We propose that it is the change in shape of the ICL mesh unit that causes the basal body misalignment. The basal body that should be safely attached to the ICL in a corner of a mesh unit is no longer anchored in place and rotates out of alignment with its rootlets attached at normal angles; the number of basal bodies per mesh unit also greatly deviates from normal, adding to the chaos of the basal body misalignment (Fig.4.2).

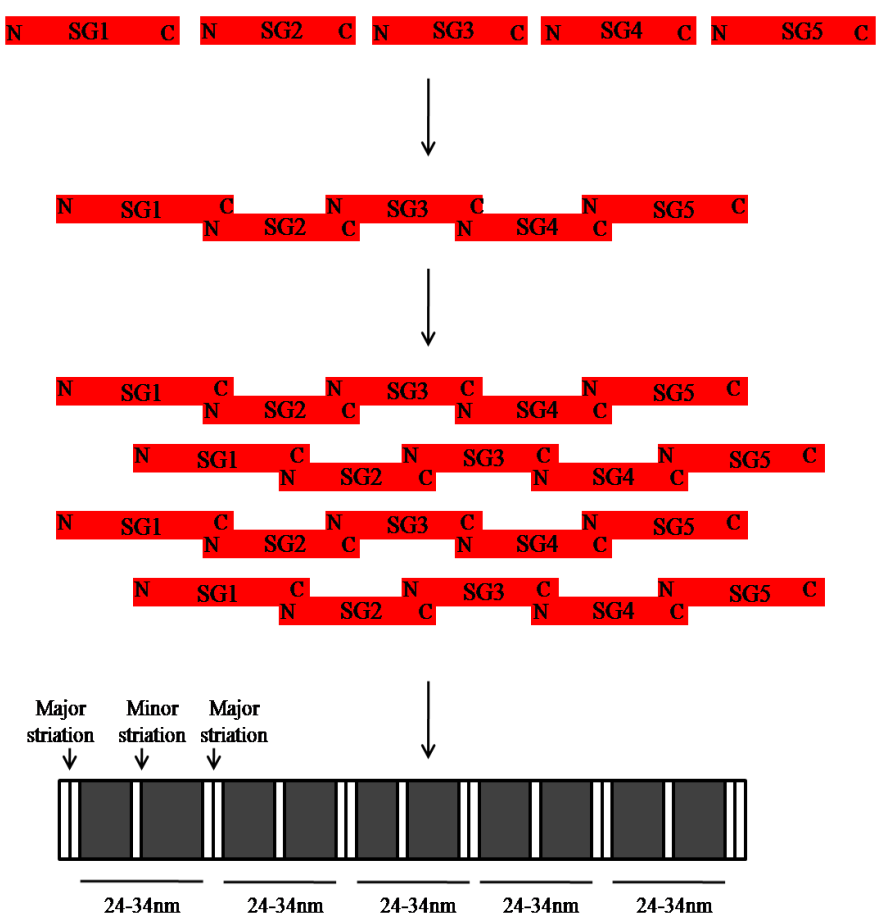
In the mouse, model depletion of rootletin affects the long-term stability of some cells (e.g., photoreceptor cells) and the function of multiciliated cells (e.g., mucus clearance activity of the lung epithelial cells). The data presented in this study provide a basis to investigate whether a reciprocal relationship exists in the higher mammalian system between the SR of the basal body and centrin related proteins that may interact to provide the long-term stability of the basal body/cilium structure or to dictate the cell surface morphology in cells with a primary or multiple cilia.

In another follow up project, we investigated the role of MKS5 protein in *Paramecium*. *Paramecium* has proven to be an excellent model to study the role of MKS3 in previous study (Picariello et al. 2014). The highly organized basal body rows at the cell have made it possible to discover basal body polarity defects, which would have been difficult to uncover in other ciliary systems. Building off the MSK3 study, we silenced another transition zone protein MKS5 using RNAi. Our immunofluorescence data demonstrates MKS5 silencing cause ciliary loss, but unlike MKS3 does not affect the basal body rows or the polarity of the basal bodies. Strikingly our data suggest

silencing of MKS5 also affects the localization of another transition zone protein B9D2 in *Paramecium*. The composition of the *Paramecium* transition zone is not fully identified yet. Our work provides a sound basis to tease apart the composition of transition zone in *Paramecium* and may yield insight the function of transition zone protein or may lead to the identification of transition zone proteins.

Fig. 4.1: Model for the assembly of the 5 different Structural Groups to constitute SRs in *Paramecium tetraurelia*.

Monomers of 5 different Structural Groups (SG1~34nm), (SG2~31nm), (SG~32nm), (SG4~30nm) and (SG5~31nm) interact to form pentamer thorough the ionic interactions of the coiled coil domains of the monomers. Order of the SR Structural Groups in the pentamer is arbitrary. Pentamers then are staggered in protofilaments. Finally, bundling of protofilaments results in the assembly of full length SR. This arrangements result in the formation of major and minor striations. The periodicity of the major striation is 24-34nm.



Monomers of 5 different Structural Group (SG)
N denotes the N-terminal of monomer
N denotes the C-terminal of monomer

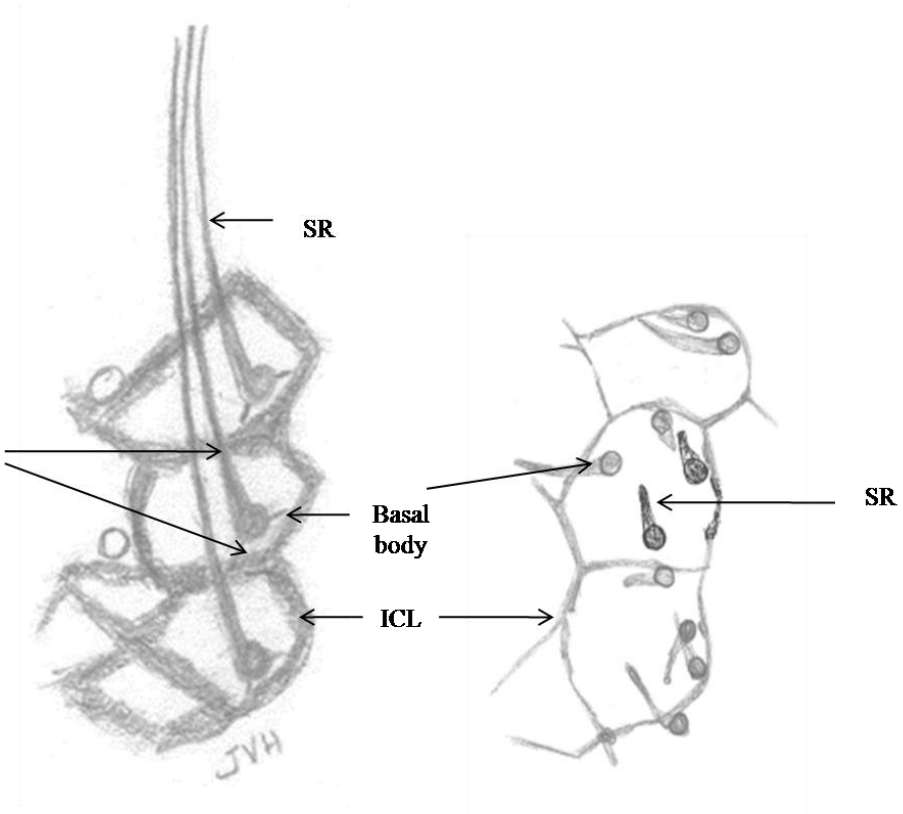
Different Structural Groups interact to form pentamer by overlapping the N-terminal and C-terminal of the rod domain of the monomers

Protofilaments formation by staggering of different SR proteins

Bundling of protofilaments results in the formation of characteristics major and minor striation patterns in assembled SR

Fig. 4.2: A reciprocal relationship exists between the cytoskeleton structures SR and ICL.

Both the SR and ICL function together to main the highly organized cortex structure in *Paramecium tetraurelia*. The diagrams in this figure are based on the images from ICL1e depletion and controls. Normally, the basal body interacts with the ICL at a corner of a mesh unit, and the SR emanating from the basal body interacts with another point on the mesh unit to act as a scaffold or strap to keep the mesh unit supported. When the ICL is depleted of ICL1e, there are many basal bodies in one mesh unit and the SR cannot make contact with the correct part of the mesh unit to participate in maintaining the mesh unit shape. Note that the SRs are shorter and misoriented when they cannot interact with ICL1e.



Control cell

SR Structural Group 2 Depleted cell

References:

Abdelhamed, Z.A., S. Natarajan, G. Wheway, C. F. Inglehearm, C. Toomes, C. A. Johnson, and D. J. Jagger. 2015. The Meckel-Gruber syndrome protein TMEM67 controls basal body positioning and epithelial branching morphogenesis in mice via the non-canonical Wnt pathway. Dis Model Mech. 8(6):527-541.

Adams, M., R. Simms, Z. Abdelhamed, H. Dawe, K. Szymanska, C. Logan, G. Wheway, E. Pitt, K. Gull, M. Knowles, E. Blair, S. Cross, J. Sayer, and C. Johnson. 2012. A meckelin-filamin A interaction mediates ciliogenesis. Hum Mol Genet. 21:1272 - 1286.

Afzelius, B. A. (2004). Cilia-related diseases. J. Pathol.204, 470-477.

Agrawal, N., P.V.N. Dasaradhi, A. Mohmmed, P. Malhotra, R.K. Bhatnagar, and S.K. Mukherjee. 2003. RNA interference: biology, mechanism, and applications. Microbiol Mol Biol Rev. 67:657-685.

Allen, R. 1971. Fine structure of membranous and microfibrillar systems in the cortex of *Paramecium caudatum*. J Cell Biol. 49:1 - 20.

Allen, R., and A. Fok. 1980. Membrane recycling and endocytosis in *Paramecium* confirmed by horseradish peroxidase pulse-chase studies. J Cell Sci. 45:131 - 145.

Arnaiz, O. and L. Sperling. 2011. ParameciumDB in 2011: new tools and new data for functional and comparative genomics of the model ciliate *Paramecium tetraurelia.* Nucleic Acids Res 39: Database issue D632-6. PMID: 20952411

Aubusson-Fleury, A., M. Lemullois, N. Garreau de Loubresse, C. Laligne, J. Cohen, O. Rosnet, M. Jerka-Dziadosz, J. Beisson, and F. Koll. 2012. The conserved centrosomal protein FOR20 is required for assembly of the transition zone and basal body docking at the cell surface. J Cell Sci. 125:4395 - 4404.

Aubusson-Fleury, A., G. Bricheux, R. Damaj, M. Lemullois, G. Coffe, F. Donnadieu, F. Koll, B. Vigues, and P. Bouchard. 2013. Epiplasmins and Epiplasm in *Paramecium*: The Building of a Submembranous Cytoskeleton. Protist. 164:451-469.

Aubusson-Fleury, A., G. Balavoine, M. Lemullois, K. Bouhouche, J. Beisson, F. Koll. 2017. Centrin diversity and basal body patterning across evolution: new insights from *Paramecium*. Biology Open 2017, 6: 765-776.

Aury, J.-M., O. Jaillon, L. Duret, B. Noel, C. Jubin, B.M. Porcel, B. Segurens, V. Daubin, V. Anthouard, N. Aiach, O. Arnaiz, A. Billaut, J. Beisson, I. Blanc, K. Bouhouche, F. Camara, S. Duharcourt, R. Guigo, D. Gogendeau, M. Katinka, A.-M. Keller, R. Kissmehl, C. Klotz, F. Koll, A. Le Mouel, G. Lepere, S. Malinsky, M. Nowacki, J.K. Nowak, H. Plattner, J. Poulain, F. Ruiz, V. Serrano, M. Zagulski, P. Dessen, M. Betermier, J. Weissenbach, C. Scarpelli, V. Schachter, L. Sperling, E. Meyer,

J. Cohen, and P. Wincker. 2006. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. Nature. 444:171-178.

Badano, J. L., T. M. Teslovich, and N. Katsanis. 2005. The centrosome in human genetic disease.Nat. Rev. Genet.6, 194-205.

Badano, J. L., N. Mitsuma, P. L. Beales, and N. Katsanis. 2006. The Ciliopathies: An Emerging Class of Human Genetic Disorders. Annu. Rev. Genomics Hum. Genet. 7, 125-148.

Beisson, J., Clerot, J. C., Fleury-Aubusson, A., Garreau de Loubresse, N., Ruiz, F. and Klotz, C. 2001. Basal body-associated nucleation center for the centrin based cortical cytoskeletal network in *Paramecium*. *Protist* 152, 339-354.

Berger, J.D. 1986. Autogamy in Paramecium. Cell cycle stage-specific commitment to meiosis. Exp Cell Res. 166:475-485.

Blacque, O., and M. Leroux. 2006. Bardet-Biedl syndrome: an emerging pathomechanism of intracellular transport. Cellular and molecular life sciences: CMLS. 63:2145 - 2161.

Blacque, O.E., C. Li, P.N. Inglis, M.A. Esmail, G. Ou, A.K. Mah, D.L. Baillie, J.M. Scholey, and M.R. Leroux. 2006. The WD repeat-containing protein IFTA-1 is required for retrograde intraflagellar transport. Mol Biol Cell. 17:5053-5062.

Bloodgood, R. 2010. Sensory reception is an attribute of both primary cilia and motile cilia. Journal of cell science. 123:505 - 509.

Braddock, S.R., K.M. Henley, and B.L. Maria. 2007. The face of Joubert syndrome: a study of dysmorphology and anthropometry. Am J Med Genet A. 143A:3235-3242.

Brehm, P., K. Dunlap, and R. Eckert. 1978. Calcium-dependent repolarization in Paramecium. J Physiol. 274:639 - 654.

Carvalho-Santos, Z., J. Azimzadeh, J.B. Pereira-Leal, and M. Bettencourt-Dias. 2011. Tracing the origins of centrioles, cilia, and flagella. The Journal of Cell Biology. 195:341.

Christensen, S., L. Pedersen, L. Schneider, and P. Satir. 2007. Sensory cilia and integration of signal transduction in human health and disease. Traffic. 8:97 - 109.

Cole, D.G., D.R. Diener, A.L. Himelblau, P.L. Beech, J.C. Fuster, and J.L. Rosenbaum. 1998. Chlamydomonas kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in Caenorhabditis elegans sensory neurons. The Journal of cell biology. 141:993-1008.

Collins, T., R.L. Baker, J.M. Wilhelm, and J.B. Olmsted. 1980. A cortical scaffold in the ciliate Tetrahymena. J Ultrastruct Res. 70:92-103.

Czarnecki, P., and J. Shah. 2012. The ciliary transition zone: from morphology and molecules to medicine. Trends Cell Biol. 22:201 - 210.

Cui, C., B. Chatterjee, D. Francis, Q.Yu, J. T. SanAgustin, R. Francis, T. Tansey, C.Henry, B. Wang, and B. Lemley. 2011. Disruption of Mks1 localization to the mother centriole causes cilia defects and developmental malformations in Meckel-Gruber syndrome. Dis Model Mech 4, 43–56.

Damaj R., S. Pomel, G. Bricheux, G. Coffe, B. Viguès, V. Ravet, P. Bouchard. 2009. Cross-study analysis of genomic data defines the ciliate multigenic epiplasmin family: strategies for functional analysis in *Paramecium tetraurelia*. BMC Evol Biol 9:125

Davis, E.E., M. Brueckner, and N. Katsanis. 2006. The Emerging Complexity of the Vertebrate Cilium: New Functional Roles for an Ancient Organelle. Developmental Cell. 11:9-19.

Dawe, H., U. Smith, A. Cullinane, D. Gerrelli, P. Cox, J. Badano, S. Blair-Reid, N. Sriram, N. Katsanis, T. Attie-Bitach, S. Afford, A. Copp, D. Kelly, K. Gull, and C. Johnson. 2007. The Meckel-Gruber syndrome proteins MKS1 and meckelin interact and are required for primary cilium formation. Hum Mol Genet. 16:173 - 186.

Dawe, H., M. Adams, G. Wheway, K. Szymanska, C. Logan, A. Noegel, K. Gull, and C. Johnson. 2009. Nesprin-2 interacts with meckelin and mediates ciliogenesis via remodelling of the actin cytoskeleton. J Cell Sci. 122:2716 - 2726.

Dippell, R. V. 1955. A temporary stain for Paramecium and other ciliate protozoa. Stain Technol. 30:69-71.

Dryl, S. 1974. Behavior and Motor response of *Paramecium*, in Paramecium, A current survey, ed. W. J. V. Wagtendonk, Elsevier, Amsterdam, London, New York.

Dykxhoorn, D.M., and J. Lieberman. 2005. The silent revolution: RNA interference as basic biology, research tool, and therapeutic. Annu Rev Med. 56:401-423.

Eckert, R. 1972. Bioelectric control of ciliary activity. Science. 176:473 - 481.

Emmer, B. T., D. Maric, and D. M. Engman. 2010. Molecular mechanisms of protein and lipid targeting to ciliary membranes. J. Cell Sci. 123:529–536.

Fliegauf, M., T. Benzing, and H. Omran. 2007. When cilia go bad: cilia defects and ciliopathies. Nat. Rev. Mol. Cell Biol.8:880–893.

Finetti, F., S.R. Paccani, J. Rosenbaum, and C.T. Baldari, 2011. Intraflagellar transport: a new player at the immune synapse. Trends Immunol. 32(4): 139–145.

Fire, A., S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, and C.C. Mello. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 391:806-811.

Follit, J.A., F. Xu, B.T. Keady, and G.J. Pazour. 2009. Characterization of mouse IFT complex B. Cell motility and the cytoskeleton. 66:457-468.

Frankel, J. 2000. Cell biology of Tetrahymena thermophila. Methods Cell Biol. 62:27–125.

Galati, D.F., S. Bonney, Z. Kronenberg, C. Clarissa, M. Yandell, N.C. Elde, M. Jerka-Dziadosz, T.H. Giddings, J. Frankel, and C.G. Pearson. 2014. DisAp-dependent striated fiber elongation is required to organize ciliary arrays. J. Cell Biol. 207, 705–715.

Galvani, A., and L. Sperling. 2002. RNA interference by feeding in Paramecium. Trends Genet. 18:11 - 12.

Garcia-Gonzalo, F., K. Corbit, M. Sirerol-Piquer, G. Ramaswami, E. Otto, T. Noriega, A. Seol, J. Robinson, C. Bennett, D. Josifova, J. Garcia-Verdugo, N. Katsanis, F. Hildebrandt, and J. Reiter. 2011. A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition. Nat Genet. 43:776 - 784.

Garreau de Loubresse, N., G., Keryer, Vigues, B. and and Beisson, J. 1988. A contractile cytoskeletal network of *Paramecium*: the infraciliary lattice. J. Cell Sci. 90:351-336.

Garreau de Loubresse, N., Klotz, C., Vigues, B., Rutin, B. and Beisson, J. 1991. Ca2+-binding proteins and contractility of the infraciliary lattice in *Paramecium*. Biol. Cell. 71: 217-225.

Gherman, A. E. E. Davis, and N. Katsanis. 2006. The ciliary proteome database: an integrated community resource for the genetic and functional dissection of cilia. Nat Genet 38:961-962.

Godiska, R., K.J. Aufderheide, D. Gilley, P. Hendrie, T. Fitzwater, L.B. Preer, B. Polisky, and J.R. Preer, Jr. 1987. Transformation of *Paramecium* by microinjection of a cloned serotype gene. Proceedings of the National Academy of Sciences of the United States of America. 84:7590-7594.

Gogendeau, D., Beisson, J., Garreau de Loubresse, N., Le Caer, J.-P., Ruiz, F., Cohen, J., Sperling, L., Koll, F. and Klotz, C. 2007. A Sfi1p-like centrin-binding protein mediates centrin based Ca2+-dependent contractility in *Paramecium*. Eukaryot. Cell. 6:1992-2000.

Gogendeau, D., Klotz, C., Arnaiz, O., Malinowska, A., Dadlez, M., Garreau de Loubresse, N., Ruiz, F., Koll, F. and Beisson, J. 2008. Functional diversification of centrins and cell morphological complexity. J. Cell Sci. 121, 65-74.

Hao, L., and J.M. Scholey. 2009. Intraflagellar transport at a glance. Journal of cell science. 122:889-892.

Heins, S. and Aebi, U. (1994). Making heads and tails of intermediate filament assembly, dynamics and networks. Curr. Opin. Cell Biol. 6: 25-33.

Hyams, J . S ., and G. G . Borisy . 1975 . Flagellar coordination in *Chlamydomonas reinhardtii* : isolation and reactivation of the flagellar apparatus .Science (Wash . DC). 189:891-893.

Hyams, J., and C. King. 1985. Identification of proteins of the striated rootlet of *Tetrahymena* by immunofluorescence microscopy and immunoblotting with an anti-rootlet serum. European Journal of Cell Biology. 38:102-105.

Huang, K., D.R. Diener, A. Mitchell, G.J. Pazour, G.B. Witman, and J.L. Rosenbaum. 2007. Function and dynamics of PKD2 in *Chlamydomonas reinhardtii* flagella. The Journal of cell biology. 179:501-514.

Hufnagel, L., A . 1969 . Cortical ultrastructure of *Paramecium aurelia*. Studies on isolated pellicles. J. Cell Biol. 40 :779 .

Iftode, F., J. Cohen, F. Ruiz, A. Torres Rueda, L. Chen-Shan, A. Adoutte, and J. Beisson. 1989. Development of surface pattern during division in *Paramecium*. I. Mapping of duplication and reorganization of cortical cytoskeletal structures in the wild type. Development. 105:191 - 211.

Iftode, F., A. Adoutte, and A. Fleury. 1996. The surface pattern of *Paramecium tetraurelia* in interphase: an electron microscopic study of basal body variability, connections with associated ribbons and their epiplasmic environment. Eur J Protistol. 32:46 - 57.

Iftode, F., and A. Fleury-Aubusson. 2003. Structural inheritance in *Paramecium*: ultrastructural evidence for basal body and associated rootlets polarity transmission through binary fission. Biol Cell. 95:39 - 51.

Jana, S.C., G. Marteil, and M. Bettencourt-Dias. 2014. Mapping molecules to structure: unveiling secrets of centriole and cilia assembly with near-atomic resolution. Curr. Opin. Cell Biol. 26:96–106.

Jerka-Dziadosz, M. 1990. Domain-specific changes of ciliary striated rootlets during the cell cycle in the hypotrich ciliate *Paraurostyla weissei*. J Cell Sci 96: 617–630.

Jerka-Dziadosz, M., F. Koll, D. Włoga, D. Gogendeau, N. Garreau de Loubresse, F. Ruiz, S. Fabczak, and J. Beisson. 2013. A Centrin3- dependent, transient, appendage of the mother basal body guides the positioning of the daughter basal body in *Paramecium*. Protist 164: 352-368.

Katsanis, N., J.R. Lupski, and P.L. Beales. 2001. Exploring the molecular basis of Bardet-Biedl syndrome. Hum Mol Genet. 10:2293-2299.

Ketting, R.F., and R.H.A. Plasterk. 2004. What's new about RNAi? Meeting on siRNAs and miRNAs. EMBO Rep. 5:762-765.

Kilburn, C. L., C. G. Pearson, E. P. Romijn, J. B. Meehl, T. H. Giddings, B. P. Jr, Culver, J. R. Yates, and M. Winey. 2007. New *Tetrahymena* basal body protein components identify basal body domain structure. J. Cell Biol. 178: 905-912.

Kobayashi, T. and Dynlacht, B.D. 2011. Regulating the transition from centriole to basal body. J. Cell Biol. 193:435–444.

Kozminski, K.G., K.A. Johnson, P. Forscher, and J.L. Rosembaum. 1993. A motility in the eukaryotic flagellum unrelated to falgellar beating. Proc Natl Acad Sci U S A. 90:5519-5523.

Kung, C., S.Y. Chang, Y. Satow, J.V. Houten, H. Hansma. 1975. Genetic dissection of behavior in *Paramecium*. Science. 188:898–904.

Kyttala, M., Tallila, J., Salonen, R., Kopra, O., Kohlschmidt, N., Paavola-Sakki, P., Peltonen, L. and Kestila, M. (2006). MKS1, encoding a component of the flagellar apparatus basal body proteome, is mutated in Meckel syndrome. Nat Genet 38(2):155–157.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London). 227:680-685.

Larson, D. E., and A. D. Dingle. 1981. Isolation, ultrastructure and protein composition of the flagellar rootlet of *Naegleria gruberi* . J. Cell Biol. 89:424-432.

Lauga, E. and Powers T. R. 2009. The hydrodynamics of swimming microorganisms, Rep. Prog. Phys. 72: 096601.

Lechtrek, K.-F., and M. Melkonian. 1991. Striated Microtubule-associated Fibers: Identification of Assemblin, A Novel 34-kD Protein that forms Paracrystals of 2-nm Filaments In Vitro. Journal of Cell Biology. 115:705-716.

Lechtrek, K.-F., and M. Melkonian. 1998. SF-Assemblin, Striated Fibers, and Segmented Coiled Coil Proteins. Cell Motility and the Cytoskeleton. 41:289-296.

Lechtreck, K.F., Rostmann, J. and Grunow, A. 2002. Analysis of *Chlamydomonas* SF-assemblin by GFP-tagging and expression of antisense constructs. J. Cell Sci. 115:1511–1522.

Lemullois, M., Klotz, C. and Sandoz, D. 1991. Evolutionary conservation of an epitope associated with striated rootlets in different epithelial ciliated cells. Biology of the Cell. 71: 201–208.

Letunic, I., and P. Bork. 2017. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2017.

Liu, J., M.A. Carmell, F.V. Rivas, C.G. Marsden, J.M. Thomson, J.-J. Song, S.M. Hammond, L. Joshua-Tor, and G.J. Hannon. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. Science (New York, N Y). 305:1437-1441.

Lodh, S., J. Yano, M. S. Valentine, and J. L. Van Houten. 2016. Voltage-gated calcium channels of *Paramecium* cilia. J. Exp. Biol. 219:3028–3038.

Lucker, B.F., M.S. Miller, S.A. Dziedzic, P.T. Blackmarr, and D.G. Cole. 2010. Direct interactions of intraflagellar transport complex B proteins IFT88, IFT52, and IFT46. J Biol Chem. 285:21508-21518.

Lupas, A., M. Van Dyke, and J. Stock. 1991. Predicting Coiled Coils from Protein Sequences, Science. 252:1162-1164.

Mason, J.M., and K.M. Arndt. 2004. Coiled coil domains: stability, specificity, and biological implications. Chembiochem. 2005(2):170-176.

Morga, B., and Bastin, P. 2013. Getting to the heart of intraflagellar transport using *Trypanosoma* and *Chlamydomonas* models: the strength is in their differences. Cilia 2:16.

Nachury, M., A. Loktev, Q. Zhang, C. Westlake, J. Peranen, A. Merdes, D. Slusarski, R. Scheller, J. Bazan, and V. Sheffield. 2007. A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. Cell. 129:1201 - 1213.

Nahon, P., Coffe, G., Le Guyader, H., Darmanaden-Delorme, J., Jeanmaire-Wolfe, R., Clerot, J.-C., and Adoutte, A. (1993). Identification of the epiplasmins, a new set of cortical proteins of the membrane cytoskeleton in *Paramecium.* J. Cell Sci. 104, 975-990.

Nakamura, T., N. Mine, E. Nakaguchi, A. Mochizuki, M. Yamamoto, K. Yashiro, C. Meno, and H. Hamada. 2006. Generation of robust left-right asymmetry in the mouse embryo requires a self-enhancement and lateralinhibition system. Dev. Cell. 11: 495-504.

Patel, H., K.-F. Lechtrek, M. Melkonian, and E. Mandelkow. 1992. Structure of Striated Microtubule-associated Fibers of Flagellar Roots. Journal of Molecular Biology. 227:698-710.

Pazour, G.J., A. Koutoulis, S.E. Benashski, B.L. Dickert, H. Sheng, R.S. Patel-King, S.M. King, and G.B. Witman. 1999. LC2, the chlamydomonas homologue of the t complex-encoded protein Tctex2, is essential for outer dynein arm assembly. Mol Biol Cell. 10:3507-3520.

Pazour, G., B. Dickert, Y. Vucica, E. Seeley, J. Rosenbaum, G. Witman, and D. Cole. 2000. Chlamydomonas IFT88 and its mouse homologue, polycystic kidney disease gene Tg737, are required for assembly of cilia and flagella. J Cell Biol. 151:709 - 718.

Pazour, G.J., S.A. Baker, J.A. Deane, D.G. Cole, B.L. Dickert, J.L. Rosenbaum, G.B. Witman, and J.C. Besharse. 2002. The intraflagellar transport protein, IFT88, is essential for vertebrate photoreceptor assembly and maintenance. The Journal of cell biology. 157:103-113.

Pedersen, L.B., and J.L. Rosenbaum. 2008. Intraflagellar transport (IFT) role in ciliary assembly, resorption and signalling. Curr Top Dev Biol. 85:23-61.

Picariello, T., M. Valentine, J. Yano, and J. Van Houten. 2014. Reduction of meckelin leads to general loss of cilia, ciliary microtubule misalignment and distorted cell surface organization. Cilia. 3:2.

Plattner, H., and R. Kissmehl. 2003. Molecular aspects of membrane trafficking in *Paramecium.* Int Rev Cytol. 232:185 - 216.

Rayner, C.F., A. Rutman, A. Dewar, M.A. Greenstone, P.J. Cole, and R. Wilson. 1996. Ciliary disorientation alone as a cause of primary ciliary dyskinesia syndrome. Am. J. Respir. Crit. Care Med. 153:1123–1129.

Ringo, D.L. 1967. Flagellar motion and fine structure of the flagellar apparatus in *Chlamydomonas*. The Journal of cell biology. 33:543-571.

Robert, A., G. Margall-Ducos, J.-E. Guidotti, O. Brégerie, C. Celati, C. Bréchot, and C. Desdouets. 2007. The intraflagellar transport component IFT88/polaris is a centrosomal protein regulating G1-S transition in non-ciliated cells. Journal of Cell Science. 120:628-637.

Rohatgi, R. and W.J. Snell. 2010. The ciliary membrane. Curr. Opin. Cell Biol. 22:541–546.

Rosenbaum, J.L., and G.B. Witman. 2002. Intraflagellar transport. Nat Rev Mol Cell Biol. 3:813-825.

Rubin, R. W., and Cunningham, W. P. 1973 . Partial purification and phosphotungstate solubilization of basal bodies and kinetodesmal fibres from *Tetrahymena pyriformis*. J. Cell Biol. 57:601-612.

Salonen, R., and R. Norio. 1984. The Meckel syndrome in Finland: epidemiologic and genetic aspects. Am J Med Genet. 18:691 - 698.

Salonen, R., and P. Paavola. 1998. Meckel syndrome. J Med Genet. 35:497 - 501.

Sang, L., J.J. Miller, K.C. Corbit, R.H. Giles, M.J. Brauer, E.A. Otto, L.M. Baye, X. Wen, S.J. Scales, M. Kwong, E.G. Huntzicker, M.K. Sfakianos, W. Sandoval, J.F.

Bazan, P. Kulkarni, F.R. Garcia-Gonzalo, A.D. Seol, J.F. O'Toole, S. Held, H.M. Reutter, W.S. Lane, M.A. Rafiq, A. Noor, M. Ansar, A.R.R. Devi, V.C. Sheffield, D.C. Slusarski, J.B. Vincent, D.A. Doherty, F. Hildebrandt, J.F. Reiter, and P.K. Jackson. 2011. Mapping the NPHP-JBTS-MKS protein network reveals ciliopathy disease genes and pathways. Cell. 145:513-528.

Saitou, N., and M. Nei.1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

Sasner, J.M.a.V.H., J. L. 1989. Evidence for a *Paramecium* folate chemoreceptor. Chemical Senses. 14:587-595.

Satir, P., T. Heuser, W. Sale. 2014. A structural basis for how motile cilia beat. *BioScience*. *64*:1073–1083.

Satow, Y., and C. Kung. 1980. Ca-induced K+outward current in Paramecium tetraurelia. J Exp Biol. 88:293 - 303.

Sharma, N., N.F. Berbari, and B.K. Yoder. 2008. Ciliary dysfunction in developmental abnormalities and diseases. Curr Top Dev Biol. 85:371-427.

Sleigh, M . A ., and N . R . Silvester. 1983 . Anchorage functions of the basal apparatus of cilia. J. Submicrosc. Cytol. 15 :101-104 .

Sloboda, R. D. (2005). Intraflagellar transport and the flagellar tip complex. J. Cell Biochem. 94, 266-272.

Smith, U., M. Consugar, L. Tee, B. McKee, E. Maina, S. Whelan, N. Morgan, E. Goranson, P. Gissen, S. Lilliquist, I. Aligianis, C. Ward, S. Pasha, R. Punyashthiti, S. Sharif, P. Batman, C. Bennett, C. Woods, C. McKeown, M. Bucourt, C. Miller, P. Cox, L. Al-Gazali, R. Trembath, V. Torres, T. Attie-Bitach, D. Kelly, E. Maher, V. Gattone, P. Harris, and C. Johnson. 2006. The transmembrane protein meckelin (MKS3) is mutated in Meckel-Gruber syndrome and the wpk rat. Nat Genet. 38:191 - 196.

Sperling, L. 1989. Isolation and partial characterization of ciliary rootlets from *Paramecium tetraurelia*. *In* Cytoskeletal and Extracellular Proteins. P.D.U. Aebi, and P.D.J. Engel, editors. Springer Berlin, Heidelberg. 316–318.

Sperling, L., G. Keryer, F. Ruiz, and J. Beisson. 1991. Cortical morphogenesis in *Paramecium*: a transcellular wave of protein phosphorylation involved in ciliary rootlet disassembly. *Dev. Biol.* 148:205–218.

Stewart, M. 1993. Intermediate filament structure and assembly. Cur. Opin. Cell Biol. 5:3-11.

Szymanska, K., and C.A. Johnson. 2012. The transition zone: an essential functional compartment of cilia. Cilia. 1:10.

Tallila, J., E. Jakkula, L. Peltonen, R. Salonen, and M. Kestila. 2008. Identification of CC2D2A as a Meckel syndrome gene adds an important piece to the ciliopathy puzzle. Am J Hum Genet. 82 (6):1361–136.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30: 2725-2729.

Tassin A.M., M. Lemullois, and A. Aubusson-Fleury. 2016. *Paramecium tetraurelia* basal body structure. Cilia. 5: 6.

Truebestein, L., and T.A. Leonard. 2016. Coiled-coils: the long and short of it. BioEssays. 38:903–916.

Tsao, C.-C., and M.A. Gorovsky. 2008. Tetrahymena IFT122A is not essential for cilia assembly but plays a role in returning IFT proteins from the ciliary tip to the cell body. Journal of cell science. 121:428-436.

Valente, E.M., C.V. Logan, S. Mougou-Zerelli, J.H. Lee, J.L. Silhavy, F. Brancati, M. Iannicelli, L. Travaglini, S. Romani, B. Illi, M. Adams, K. Szymanska, A. Mazzotta, J.E. Lee, J.C. Tolentino, D. Swistun, C.D. Salpietro, C. Fede, S. Gabriel, C. Russ, K. Cibulskis, C. Sougnez, F. Hildebrandt, E.A. Otto, S. Held, B.H. Diplas, E.E. Davis, M. Mikula, C.M. Strom, B. Ben-Zeev, D. Lev, T.L. Sagie, M. Michelson, Y. Yaron, A. Krause, E. Boltshauser, N. Elkhartoufi, J. Roume, S. Shalev, A. Munnich, S. Saunier, C. Inglehearn, A. Saad, A. Alkindy, S. Thomas, M. Vekemans, B. Dallapiccola, N. Katsanis, C.A. Johnson, T. Attie-Bitach, and J.G. Gleeson. 2010. Mutations in TMEM216 perturb ciliogenesis and cause Joubert, Meckel and related syndromes. Nature genetics. 42:619-625.

Valentine, M., A. Rajendran, J. Yano, S. Weeraratne, J. Beisson, J. Cohen, F. Koll, and J. Van Houten. 2012. Paramecium BBS genes are key to presence of channels in cilia. Cilia. 1:16.

Van Houten, J. 1998. Chemosensory transduction in *Paramecium*. Eur. J. Protistol. 34: 301–307.

Vincensini, L., Blisnick, T., and Bastin, P. 2011. 1001 model organisms to study cilia and flagella. Biol Cell 103, 109-130.

Waters, A.M. and P.L. Beales. 2011. Ciliopathies: An expanding disease spectrum. Pediatr. Nephrol.26, 1039–1056.

Weber, K., N. Geisler, U. Plessmann, A. Bremerich, K.-F. Lechtrek, and M. Melkonian. 1993. SF-Assemblin, the Structural Protein of the 2-nm Filaments from Striated Microtubule Associated Fibers of Algal Flagellar Roots, Forms a Segmented Coiled Coil. J. Cell Biol. 121:837-845.

Williams, C., C. Li, K. Kida, P. Inglis, S. Mohan, L. Semenec, N. Bialas, R. Stupay, N. Chen, O. Blacque, B. Yoder, and M. Leroux. 2011. MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis. J Cell Biol. 192:1023 - 1041.

Williams, C.L., M.E. Winkelbauer, J.C. Schafer, E.J. Michaud, and B.K. Yoder. 2008. Functional redundancy of the B9 proteins and nephrocystins in Caenorhabditis elegans ciliogenesis. Mol Biol Cell. 19:2154-2168.

Yang, J., Liu, X., Yue, G., Adamian, M., Bulgakov, O., and T. Li. 2002. Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet. J. Cell Biol. 159:431–440.

Yang, J., J. Gao, M. Adamian, X.H. Wen, B. Pawlyk, L. Zhang, M.J. Sanderson, J. Zuo, C.L. Makino, and T. Li. 2005. The ciliary rootlet maintains longterm stability of sensory cilia. Mol. Cell. Biol. 25:4129–4137.

Yu, S., K. Hackmann, J. Gao, J. Gao, X. He, K. Piontek, M.A. Garcia-Gonzalez, M.A. Garcia Gonzalez, L.F. Menezes, H. Xu, G.G. Germino, J. Zuo, and F. Qian. 2007. Essential role of cleavage of Polycystin-1 at G protein-coupled receptor proteolytic site for kidney tubular structure. Proceedings of the National Academy of Sciences of the United States of America. 104:18688-18693.